Universidade de Lisboa

Faculdade de Farmácia

Rescue of cystathionine β-synthase R>C variants by thiol compounds: Accessibility of affected residues as a key player

Ana Sofia Artur Carreira de Sousa Garcia

Dissertação de Mestrado

Mestrado em Ciências Biofarmacêuticas

2015

Universidade de Lisboa

Faculdade de Farmácia

Rescue of cystathionine β-synthase R>C variants by thiol compounds: Accessibility of affected residues as a key player

Ana Sofia Artur Carreira de Sousa Garcia

Dissertação de Mestrado orientada por:

Professora Doutora Ana Paula Peralta Leandro

Professora Doutora Maria de Fátima Vieira Ventura

Mestrado em Ciências Biofarmacêuticas

2015

The studies presented in this dissertation were performed within the Department of Biochemistry and Human Biology, Metabolism and Genetics research group, at the Research Institute for Medicines (iMed.ULisboa), Faculty of Pharmacy, Universidade de Lisboa, under the supervision of Paula Leandro, Ph.D, and Fátima Ventura, Ph.D.

To the strength of being better,

not than the other,

but than the self from yesterday.

Table of Contents

Abstract xiii
Resumo xv
Abbreviations xvii
1. Introduction 1
1.1. Conformational Diseases 3
1.2. Classical Homocystinuria 4
1.3. Homocysteine and the Methionine Cycle 5
1.4. Classical Homocystinuria and Clinical Consequences 7
1.5. The CBS deficiency as a Conformational Disease 7
1.6. The CBS protein 8
1.7. The *CBS* gene mutations 9
1.7.1. R>C mutations in CBS Deficiency 10
1.8. Classical homocystinuria and therapeutic options 14
1.8.1. The rescue of p.R336C variant by thiol compounds 15
2. Aims 19
3. Material and Methods 23
3.1. Production of mutant CBS expression constructs 25
3.2. Expression of WT and Variant Proteins in *E. coli* 27
3.3. Purification of WT and Variant Proteins 28
3.4. Characterization of recombinant proteins 30
3.4.1. SDS-PAGE Analysis 30
3.4.2. Assay of Enzymatic Activity 31
3.4.3. Quantification of Cystathionine by HPLC 31
3.4.4. Limited Proteolysis 32
3.4.5. Differentiation Scanning Fluorimetry 32
3.4.6. Modification of Recombinant proteins with PEG-Mal 34
3.4.7. Evaluation of –SH groups by DTNB 34
4. Results and Discussion 37
4.1 Characterization of the R>C CBS variant proteins 39
4.1.1. Production of the R>C and R>H expression constructs 39
4.1.2. Expression and purification of recombinant WT and CBS variants 40

4.1.3. Conformational flexibility of the R>C recombinant proteins 42
4.1.4. Thermal stability of the R>C recombinant proteins 45
4.1.5. Assessment of MEG binding by PEG-Mal modifications 48
4.1.6. Enzymatic activity of R>C variants 52
4.2. Characterization of the p.L388P variant identified in a Portuguese CBS deficient patient 56
4.2.1. Expression and purification of recombinant WT and p.L388P CBS variant 57
4.6.4. Effect of PLP on the catalytic activity of the p.L338P variant 58
5. General Discussion 61
6. References 69
Agradecimentos 85

Abstract

Among the 20 amino acids incorporated into a polypeptide chain during protein synthesis, arginine (Arg; R) is the one with the highest relative mutability. In genetic diseases, almost 15% of missense mutations occur in Arg residues and more than half of these missense mutations are Arg to cysteine (Cys; C) substitutions. The Arg residue is more prone to mutations due to the fact that this residue can be coded by six different codons, four of them containing CpG dinucleotides, a preferential site for cytosine methylation with further spontaneous deamination to thymine.

The effect of the aminothiol compounds cysteamine (CySH) and mercaptoethylguanidine (MEG) in the specific rescue of R to C (R>C) mutations have already been studied in our group through the characterization of the most common variant (p.R336C) found in cystathionine β-synthase (CBS) deficiency. It was postulated that those compounds would bind to the mutant residue C, forming a structure resembling the wild-type (WT) residue R, and thus restoring the enzyme activity.

Cystathionine-β-synthase is a cytoplasmatic homotetrameric enzyme. It catalyses the condensation of the amino acids L-Homocysteine (L-Hcy) and L-Serine (L-Ser) to form L-Cystathionine (L-Cth). Each subunit (63 kDa) binds to the cofactors pyridoxal 5'-phosphate (PLP) and heme and comprises a N-terminal heme-binding domain, a catalytic core domain which binds PLP, and a C-terminal regulatory domain.

The CBS protein contains 28 Arg residues, localized all over its sequence. In order to understand if the rescue of the activity of R>C CBS variants by aminothiol compounds, such as CySH and MEG, depends on the localization of the affected residue, six residues presenting different accessible surface area (ASA) values and domain localization were selected for mutagenesis, namely the R18, R121, R164, R336, R369 and R491 residues. Additionally, a control variant corresponding to an Arg residue to histidine (His; H) substitution (R>H) was also studied.

The CBS proteins (WT and variants) were produced in a prokaryotic expression system and further purified by affinity chromatography. To evaluate the impact of aminothiol compounds on the structure and function of CBS R>C variants, the purified recombinant proteins were characterized in the absence and presence of these compounds in respect to

their thermostability by differential scanning fluorimetry (DSF), and to their conformational flexibility by susceptibility to limited proteolysis. The enzyme activity and the quantitative assessment of free Cys residues were also evaluated.

The obtained data strongly suggest that MEG, but not CySH, is able to increase the resistance to proteolysis of the majority of the studied R>C variants. However, this did not result in a higher enzyme activity for all of them. Since the observed effect was independent from ASA and domain localization, we anticipate that rescuing of R>C variants by aminothiols will be residue-specific and should always be investigated in a case-by-case basis.

Key words: Cystathionine β-synthase; Classical Homocystinuria; Conformational Diseases; Missense mutation; Aminothiol compounds; Accessible surface area; Protein rescue; Enzyme activity modulation; Cysteamine; Mercaptoethylguanidine.

Resumo

Entre os 20 aminoácidos incorporados numa cadeia polipeptídica durante a síntese proteica, a arginina (Arg; R) é o aminoácido com maior mutabilidade relativa. Nas doenças genéticas, quase 15% das mutações *missense* são substituições de arginina para cisteína (Cys; C). O resíduo Arg é altamente susceptível a mutagénese devido ao facto de este aminoácido poder ser codificado por seis codões diferentes, contendo quatro deles dinucleótidos CpG, um local preferencial para a metilação de citosina com posterior desaminação espontânea para timina.

O efeito dos compostos aminotiol cisteamina (CySH) e mercaptoetilguanidina (MEG) na recuperação específica de mutações R para C (R>C) foi já estudado no nosso grupo através da caracterização da variante mais comum (p.R336C) da deficiência na enzima cistationina β-sintase (CBS). Foi então postulado que estes compostos ligar-se-iam ao resíduo Cys mutante, formando uma estrutura semelhante ao resíduo Arg selvagem, restaurando assim a actividade enzimática.

A proteína CBS é uma enzima homotetramérica citoplasmática. Cada uma das suas subunidades (63 KDa) liga-se aos cofatores piridoxal 5'-fosfato (PLP) e heme. Esta enzima catalisa a condensação dos aminoácidos L-homocisteína (L-Hcy) e L-serina (L-Ser) para formar L-cistationina (L-Cth). Cada monómero da CBS contém três domínios: um domínio N-terminal de ligação ao heme, que inclui os primeiros 70 aminoácidos; um domínio catalítico, contendo os 340 aminoácidos centrais; e um domínio regulador C-terminal com 140 resíduos.

A CBS contém, na sua sequência primária, 28 resíduos de Arg igualmente distribuídos pelos diferentes domínios. De modo a compreender se a recuperação da atividade das variantes R>C da CBS, por compostos aminotiol como CySH e MEG, depende da localização do resíduo afetado, foram selecionados para mutagénese seis resíduos com valores de área de superfície acessível (ASA) e diferentes localizações, nomeadamente os resíduos R18, R121, R164, R336, R369 e R491. De entre estes, os resíduos R18 e R491 são resíduos expostos ao solvente e estão localizados no domínio N-terminal e C-terminal, respetivamente. O resíduo R121 apresenta um ASA de apenas 6% e localiza-se no domínio catalítico. Os resíduos R164, R336 e R369 são representativos de resíduos parcialmente

acessíveis ao solvente, localizando-se na sua totalidade no domínio catalítico. Adicionalmente, foi estudada, como controlo, uma substituição equivalente de Arg para Histidina (His; H) (R>H).

As proteínas CBS, selvagem (WT) e variantes, foram produzidas num sistema de expressão procariótico e purificadas por cromatografia de afinidade. Para avaliar o impacto dos compostos aminotiol na estrutura e função das mutações R>C da CBS, estas variantes foram caracterizadas na ausência e presença destes compostos relativamente à sua termostabilidade por fluorimetria de varrimento diferencial (DSF), e à sua suscetibilidade à proteólise limitada. A atividade enzimática e a aferição quantitativa dos resíduos de Cys livres foram também avaliadas.

Os dados obtidos sugerem que o MEG, mas não a CySH, tem capacidade para modular a flexibilidade conformacional (maior resistência à proteólise limitada) das variantes e de aumentar a atividade das variantes R>C da CBS. Uma vez que o efeito observado é independente do ASA e da localização do resíduo, antecipamos que a recuperação das variantes R>C por compostos aminotiol possa ser alargada a outras proteínas com mutações R>C, devendo no entanto este efeito sempre ser avaliado especificamente para cada variante.

Palavras-chave: Cistationina β-sintase; Homocistinúria Clássica; Doenças Conformacionais; Mutação *missense*; Compostos aminotiol; Área de superfície acessível; Recuperação proteica; Modulação da actividade enzimática; Cisteamina; Mercaptoetilguanidina.

Abbreviations

6xHis	6 histidine residues
AET	S-(2-aminoethyl)thiuronium bromide hydrobromide
Arg	Arginine
ASA	Accessible surface area
BHMT	Betaine-homocysteine methyltransferase
BSA	Bovine serum albumin
(h)CBS	(Human) Cystathionine β-synthase
CGL	Cystathionine γ-lyase
COMT	Catechol-O-methyltransferase
COX	Cyclooxygenase
L-Cth	L-Cystathionine
L-Cys	L-Cysteine
CySH	Cysteamine
DLS	Dynamic light scattering
DMSO	Dimethyl sulfoxide
DSF	Differential Scanning Fluorimetry
DTNB	5,5'-dithiobis-(2-nitrobenzoic acid)
DTT	Dithiothreitol
E. coli	*Escherichia coli*
GCL	Glutamate-cysteine ligase
GGC	γ-glutamylcysteine
GS	GSH synthase
GSH	Glutathione
L-Hcy	L-Homocysteine
His	Histidine
IMAC	Immobilized metal affinity chromatography
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
Kan	Kanamycin

LB	Luria Bertani broth
LMW	Low Molecular Weight
MAT	Methionine adenosyltransferase
MCS	Multiple cloning site
MEG	Mercaptoethylguanidine
L-Met	L-Methionine
MS	Methionine synthase
MTHFR	Methylenetetrahydrofolate reductase
NEM	N-Ethylmaleimide
ON	Overnight
PEG-Mal	Polyethylene glycol maleimide
PLP	Pyridoxal 5'-phosphate
PMSF	Phenylmethylsulfonyl fluoride
RBS	Ribosome binding site
rpm	Rotations per minute
RT	Room temperature
SAH	S-adenosylhomocysteine
SAHH	S-adenosylhomocysteine hydrolase
SAM	S-adenosylmethionine
SDS	Sodium dodecyl sulfate
PAGE	Polyacrylamide gel electrophoresis
L-Ser	L-Serine
SHMT	Serine hydroxymethyltransferase
T_{aggr}	Temperature of aggregation
THF	Tetrahydrofolate
T_m	Melting temperature
WT	Wild-type

1. Introduction

1.1. Conformational Diseases

Half of all sequence alterations in genetic diseases are missense mutations leading to single amino acid substitutions. These can affect not only the activity but also the structure, stability and folding of the resulting variant protein, causing its degradation or accumulation under the form of aggregates. These diseases are designated conformational disorders and can be categorized in loss-of-function or gain-of-function, respectively (Mendre et al., 2010).

Gain of function disorders usually occur when soluble proteins undergo conformational re-arrangements becoming capable of aggregate into β-sheets conformations, leading to the production of toxic insoluble complexes known as amyloid deposits, that accumulate extracellularly or in the cytoplasm, in the nucleus and at the cell membrane, depending on the disease (Guidolin et al., 2012).

To be biologically active, proteins must present a correct 3D arrangement, determined by the amino acid sequence. Proper folding is acquired by interactions established between specific residues, which nucleate and trigger the cooperative folding reaction into the final topology, corresponding to the most stable structure under physiological conditions. As a consequence, the misfolded protein can be targeted for degradation, leading to lower intracellular steady-state levels (loss of function disease) (Leandro et al., 2008).

Specific compounds such as substrates, inhibitors or other ligands are capable of binding to misfolded variant proteins and stabilizing them, enabling the acquisition of a near native conformation. These molecules known as pharmacological chaperones help the defective protein to escape the cell quality control system and therefore, maintain its intracellular levels (Mendre et al., 2010). The induced stabilization results essentially from a direct effect of small molecular weight (SMW) compounds that correct protein misfolding or folding defects, attempting to mimic the function of molecular chaperones (Leandro et al., 2008).

1.2. Classical Homocystinuria

Hyperhomocysteinaemia, or elevated plasma concentrations of L-homocysteine (L-Hcy) in blood, can be caused by a range of metabolic disorders, from which the most frequent is due to a deficient activity of cystathionine β-synthase (CBS; E.C. 4.2.1.22), an enzyme that condensates L-Hcy and L-serine (L-Ser) into L-cystathionine (L-Cth) (Figure 1). This condition is known as classical homocystinuria (Mudd et al., 2001) or CBS deficiency (OMIM #236200), and is caused by mutations in the *CBS* gene. Until present, more than 160 disease-associated mutations have been identified (http://cbs.lf1.cuni.cz/mutations.php).

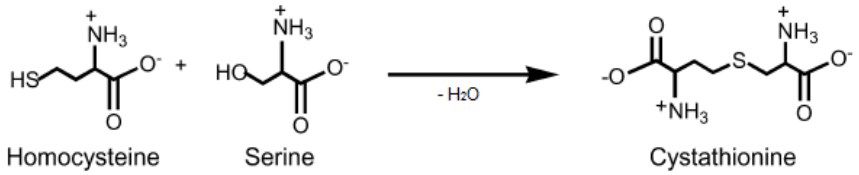

Figure 1 – Condensation reaction catalyzed by Cystathionine β-Synthase (adapted from Banerjee et al, 2005).

The *CBS* gene is localized in chromosome 21, more specifically in the short arm, 22.3 region (21q22.3) (Münke et al., 1988), and contains 23 exons codifying for a polypeptide of 551 amino acids (Figure 2).

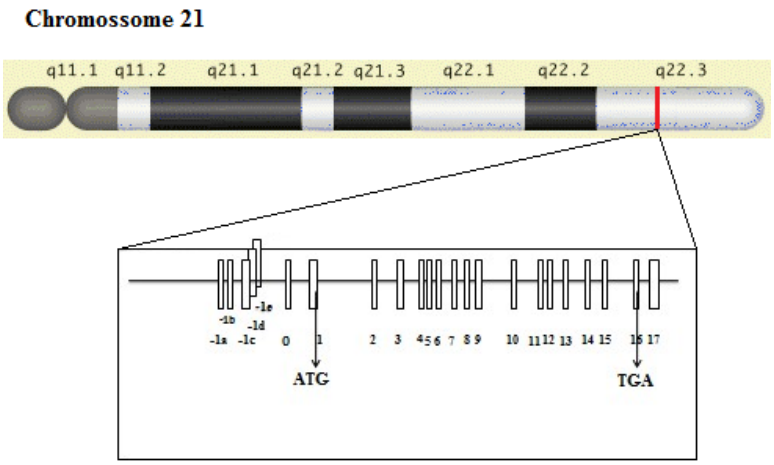

Figure 2 – Location in chromosome 21 and structure of the *CBS* gene. Exons are represented by numbered solid boxes. The beginning and the end of the coding region are indicated by the codons ATG and TGA, respectively (adapted from Kraus et al., 1998).

1.3. Homocysteine and the Methionine Cycle

Cystathionine β-synthase is an enzyme of the methionine (Met) cycle (Figure 3), which plays a key role in sulfur amino acid metabolism in mammalian cells (McKeever et al., 1991). The substrate of CBS, L-Hcy, is localized at the intersection point between the two remethylation pathways and the transsulfuration pathway (Finkelstein, 1990) (Figure 3). About half of the L-Hcy produced in the cell is conserved by remethylation as L-Met (Finkelstein et al., 1984). The other half is irreversibly converted by CBS and cystathionine γ-lyase (CGL; E.C. 4.4.1.1) to L-cysteine (L-Cys). Thus, by catalyzing the first step of the transsulfuration pathway (Figure 3), CBS is directly involved in the removal of L-Hcy from the Met cycle and in the biosynthesis of L-Cys, a precursor of glutathione, the major redox regulating metabolite of the cell (Finkelstein et al., 1984). In eukaryotes, the sulfur atom of L-Cys is derived from L-Met, while the carbon chain and the amino group originate from L-Ser (Kraus, 1998).

The methionine adenosyl transferase (MAT; E.C. 2.5.1.6) initiates the Met cycle by converting L-Met into S-adenosylmethionine (SAM) (Kraus, 1998). S-adenosylmethionine is the cellular universal donor of methyl groups yielding S-adenosylhomocysteine (SAH) and adenosine (Ado) by the activity of various methyl transferases (MT). S-adenosylhomocysteine is then converted to L-Hcy by S-adenosylhomocysteine hydrolase (SAHH; E.C. 3.3.1.1). This enzymatic step is reversible and SAH synthesis is thermodynamically favored, unless the products (Ado and L-Hcy) are removed. In one of the remethylation pathways L-Hcy is converted to L-Met by methionine synthase (MS), an enzyme that requires 5-methyltetrahydrofolate (MTHF) as the methyl donor and methylcobalamin (B12) as the coenzyme (Figure 3). The other remethylation pathway of L-Hcy to L-Met is catalyzed by a second homocysteine methylase, betaine-homocysteine methyltransferase (BHMT), and requires betaine has the methyl donor to yield N,N,dimethyl-glycine.

In contrast to the ubiquitous Met cycle, transsulfuration has a limited distribution in mammalian tissues and all cells lacking this pathway require an exogenous source of L-Cys (Finkelstein, 1990). Liver, kidney, small intestine, and pancreas contain both pathways. These tissues have a rapid turnover of glutathione, a process that consumes L-Cys (McKeever et al., 1991). Interestingly, two of the enzymes of the transsulfuration pathway (CBS and

CGL) require as a cofactor the pyridoxal 5'-phosphate (PLP), the active form of vitamin B6 or pyridoxine.

The regulation of L-Hcy or L-Met metabolism in mammalian tissues relies on the distribution of the substrate between competing reactions at two metabolic sites (Finkelstein, 1990). The reactions that utilize L-Hcy comprise one of these regulatory sites, as the methionine cycle and the transsulfuration pathway both limit the tissue concentration of L-Hcy (Finkelstein et al., 2000). The other metabolic site relates to the competition for Met between protein synthesis and the formation of SAM.

In fact, the Met cycle is tightly regulated by SAM: an elevated level of this compound activates CBS (eliminating L-Hcy in excess) and inhibits MTHFR (reducing the levels of 5-CH_3-THF, the MS substrate).

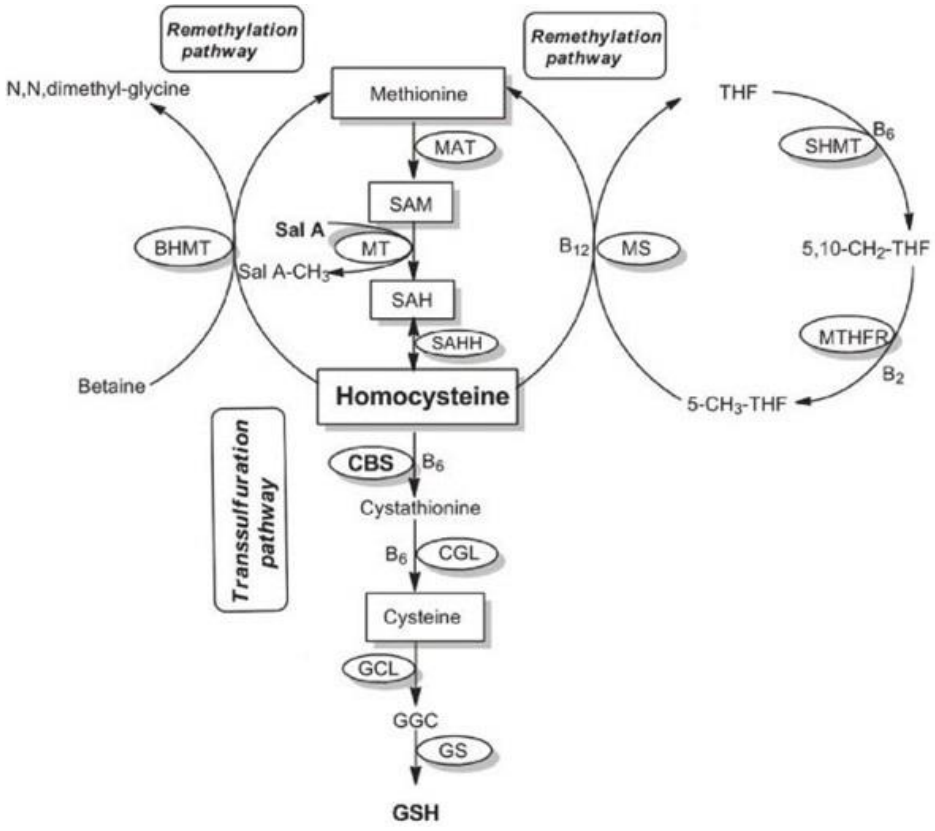

Figure 3 – Methionine cycle. MAT - Methionine adenosyltransferase; SAM - S-Adenosylmethionine; MT – Methyl transferase; SAH - S-Adenosylhomocysteine; SAHH - S-Adenosylhomocysteine hydrolase; BHMT - Betaine-homocysteine methyltransferase; MS - Methionine synthase; MTHFR - Methylenetetrahydrofolate reductase; SHMT - Serine hydroxymethyltransferase; CBS - Cystathionine β-synthase; CGL - Cystathionine γ-lyase; THF - Tetrahydrofolate; GSH – Glutathione; GGC - γ-glutamylcysteine; GCL - Glutamate-cysteine ligase; GS - GSH synthase (adapted from (Zhang et al., 2013)).

1.4. Classical Homocystinuria and Clinical Consequences

It is known that high concentrations of L-Hcy cause extensive damage to collagen and elastic fibers and may lead to a multisystemic disorder involving many organs and tissues, notably the cardiovascular and the central nervous systems, the eye, connective tissues and muscles (Mudd et al., 1985). Lens dislocation is one of the typical features presented by patients with CBS deficiency, and the most common sign leading to diagnosis (Mudd et al., 1985; Cruysberg et al., 1996). In these patients, numerous skeletal abnormalities resembling the Marfan syndrome may be observed (Mudd et al., 1989), both by clinical and X-ray examinations including scoliosis/kyphosis, dolichostenomelia (long and thin extremities), decreased upper/lower segment ratio and arachnodactyly (Skovby et al., 1999). Vascular involvement is another peculiar feature of this disease, which can be generally characterized as a thrombotic diathesis that may manifest in the venous or arterial system, and/or as an accelerated atherosclerosis (Carmel et al., 2001). Mental retardation is also a frequent finding in CBS deficient patients (Mudd et al., 1985).

Two variants of CBS deficiency are recognized: B6-responsive and B6-nonresponsive (Hu et al., 1993). Vitamin B6-responsive homocystinuria is typically (but not always) milder than the nonresponsive variant. In fact, less severe clinical phenotypes are observed when comparing untreated B6-responsive and B6-nonresponsive patients, with a lower number of B6-responsive patients presenting dislocation of lenses by the age of 10 years (55% vs. 82%), thromboembolic events by the age of 15 years (12% vs. 27%), spinal osteoporosis by the age of 15 years (36% vs. 64%), and mortality by the age of 30 years (4% vs. 23%) (Mudd et al., 2001). Quantitative data from 284 patients also showed a median IQ of 78 and 56 for the B6 responders and non-responders, respectively (Mudd et al., 1985).

1.5. The CBS deficiency as a Conformational Disease

As previously referred, the majority of disease-causing *CBS* gene mutations are missense mutations leading to single amino acid substitutions in the translated protein. Structural characterization of several of these CBS variants demonstrated that the

recombinant variants are misfolded proteins, and therefore CBS deficiency is presently recognized as a Conformational Disease (Hnízda et al., 2012).

The rescue of proper folding by chemical/pharmacological chaperones has been extensively studied in the last years as a new therapeutic approach to treat Conformational Diseases and CBS deficiency has been no exception. In fact, CBS misfolded variants have already been a target for folding rescue by the use of different compounds (Kopecká et al., 2011).

1.6. The CBS protein

Cystathionine β-synthase is a cytoplasmatic homotetrameric enzyme, and each subunit (63 kDa) binds to two cofactors (PLP and heme) and to two substrates (L- Hcy and L-Ser).

Each monomer of CBS comprises three domains: the heme-binding domain, which includes the first 70 N-terminal amino acids; the catalytic core, comprising the central 340 residues (Ala71–His411) and the C-terminal regulatory domain, with 140 residues (Aitken et al., 2005) (Figure 4). Structurally, the monomer is composed of eleven α-helices, seven short 3_{10} helices and two β-sheets consisting of four (N-terminal) and six (C-terminal) strands (Meier et al., 2001). It has been demonstrated that Lys119 is the PLP binding residue in human CBS (Kery et al., 1994). As stated before, CBS is regulated by SAM (Meier et al., 2001; Pey et al., 2013) which activates the enzyme 2-4 fold (Kožich et al., 1992). This allosteric activator binds to the C-terminal regulatory domain and recent data indicate that each monomeric subunit binds to 6 mol of SAM (Pey et al., 2013). The presence of heme in CBS is surprising and still unclear, as the mechanism of the β-replacement reactions catalyzed by the enzyme can occur and be explained solely by PLP mediated catalysis. Heme binds to CBS through an N-terminal loop that provides the axial ligands Cys52 and His65. The heme resides in a small hydrophobic pocket at the outer end of each dimer, distant from the PLP cofactor, which is deeply buried in the active site and accessible only via a narrow channel (Meier et al., 2001). The distance of heme from the PLP binding site suggests its

non-role in catalysis However, deletion of the heme domain causes loss of redox sensitivity. It is, therefore, hypothesized that the heme is a redox sensor (Yamanishi et al., 2006).

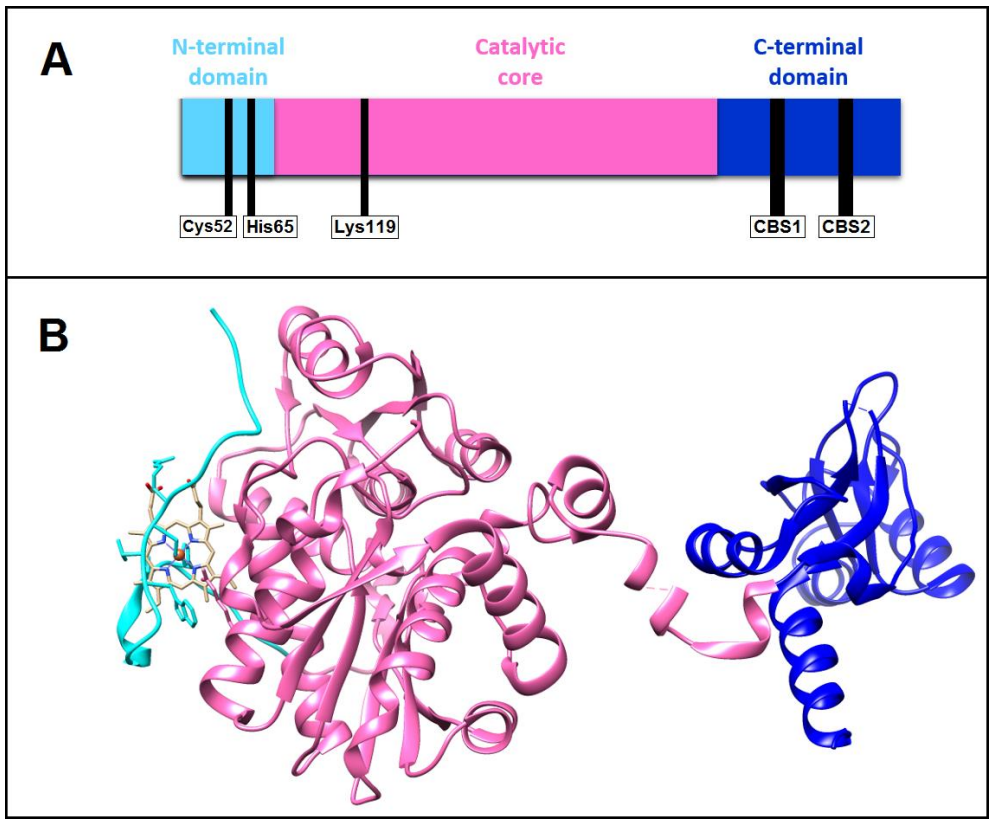

Figure 4 – Domains' structure and organization of a CBS protein monomer. (A) Schematic representation of the structural/functional domains of the full-length CBS with the N-terminal domain in cyan, the catalytic core domain in hot pink and the C-terminal regulatory domain in blue. The Cys52 and His65 residues, responsible for interaction with heme, are indicated in the N-terminal domain; the Lys119 residue, responsible for interaction with PLP, is indicated in the catalytic core domain; the CBS1 and CBS2 motifs, responsible for the interaction with SAM, are indicated in the C-terminal regulatory domain. (B) Model of the CBS monomer depicted in the same color code as in the schematic representation A. The images were obtained using the crystal structure of a truncated form of CBS (PDB ID 4L0D) and the UCSF Chimera, version 1.8.1 (build 39231)).

1.7. The *CBS* gene mutations

The incidence of CBS deficiency is thought to be 1:20,000 to 1:60,000 in most European countries (Mudd et al., 2001), but may be as low as 1/1,000,000 in some populations, like the Japanese (Mudd et al., 1989). On the opposite, an estimated incidence

of 1/6,400 has been recently reported from Norway (Refsum et al., 2004) and of 1/3,125 in Qatar. The incidence of classical homocystinuria in the Qatari population is at least 5.56 cases per 10,000 (1/1,800) representing the highest known incidence of this disease in the world (Bashir et al., 2015).

As referred above, more than 160 disease-causing mutations have been identified in the *CBS* gene (http://cbs.lf1.cuni.cz/mutations.php). The mutational spectrum differs between countries, and for many mutations it has been determined their vitamin B6 responsiveness status. Therefore, gene mutation analysis is not only of diagnostic value, but it also helps to distinguish between the two clinically relevant B6-responsive and B6-nonresponsive variants (Mudd et al., 2001).

Most of the CBS deficient patients carry missense mutations, followed by a few nonsense mutations, various deletions, insertions, and splicing mutations. About half of all point substitutions in the coding region of the *CBS* gene are originated from deaminations of methylcytosines in CpG dinucleotides (Kraus et al., 1999).

1.7.1. R>C mutations in CBS Deficiency

Among the 20 amino acids incorporated in a polypeptide chain during protein synthesis, Arginine (Arg; R) is the one with the highest relative mutability (Khan et al., 2007). In fact, in genetic diseases almost 15% of missense mutations occur in Arg residues and interestingly more than half of these missense mutations lead to Arg to Cys (R>C) substitutions (Steward et al., 2003). The Arg residue is more prone to substitutions as it can be coded by six different codons (CGU, CGC, CGA, CGG, AGA, AGG), four of them containing CpG dinucleotides (underlined), a preferential site for cytosine methylation with further spontaneous deamination to thymine (Vitkup et al., 2003; Coulondre et al., 1978).

The primary structure of the CBS protein contains 28 Arg residues (Table 1, Figure 5A), from which residues 18, 121, 336, 369 and 491 present pathogenic R>C substitutions (Figure 5B, Table 1), with a total of 93 cases (http://cbs.lf1.cuni.cz/mutations.php).

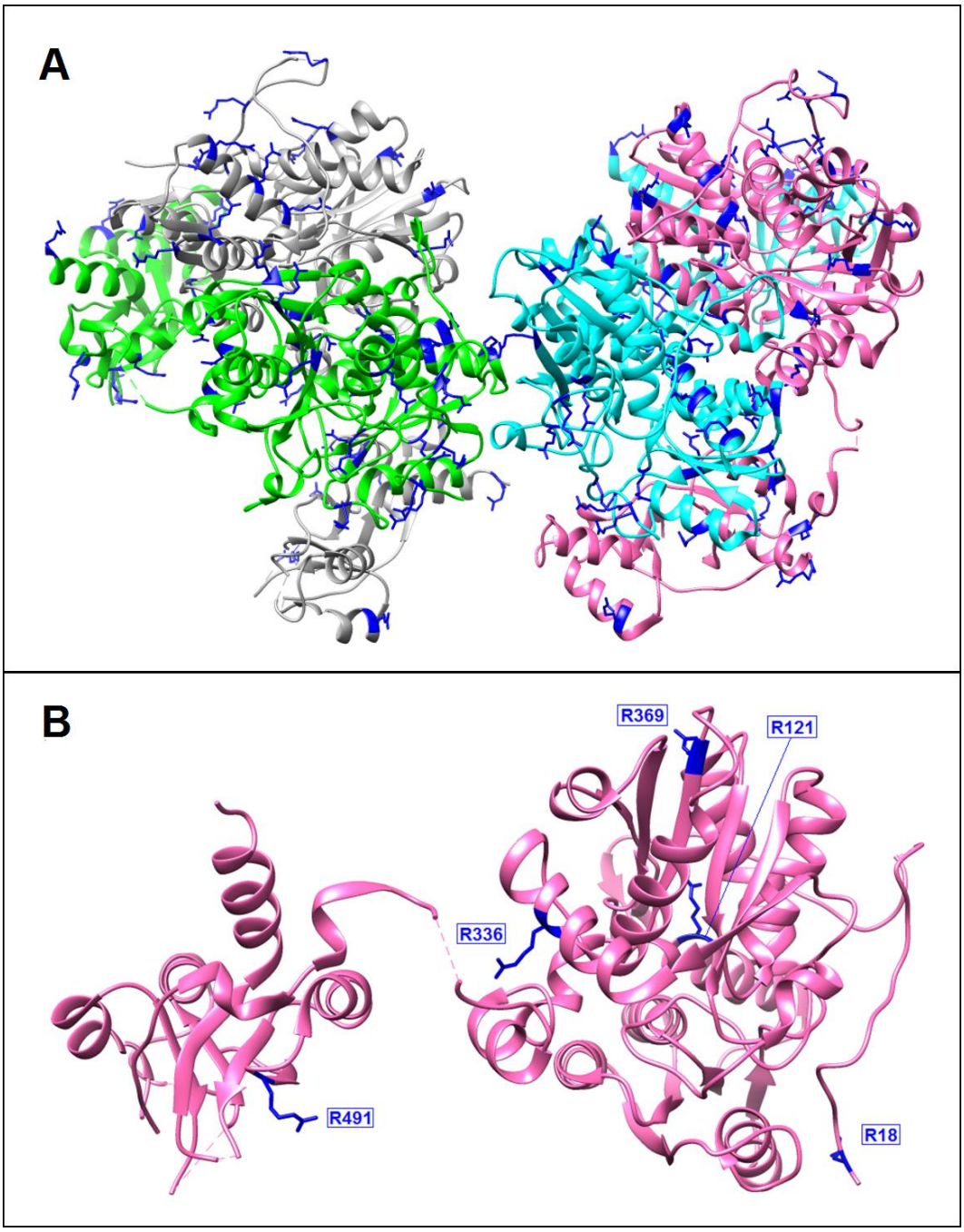

Figure 5 – Human cystathionine β-synthase (CBS). **(A)** Structural representation of the CBS tetramer with each monomer represented by a different color and Arg residues depicted in blue. **(B)** Structural representation of a CBS monomer with the Arg residues for which R>C substitutions have been described in CBS deficient patients depicted in blue. Images obtained using the crystal structure of a truncated form of CBS (PDB ID 4L27) and the UCSF Chimera, version 1.8.1 (build 39231)).

In CBS, the Arg residues are equally distributed by the three different domain (≈ 5%), being predominantly localized in α-helices (≈53.8%).

Table 1 – List of CBS Arg residues and respective characteristics

Residue	Domain localization[1]	ASA[2] (%)	SS elements[3]	Functional involvement	R>C disease-causing mutation
18	N	-	-	-	Yes
45	N	50	Random coil	-	-
51	N	99	Random coil	-	-
58	N	34	Random coil	-	-
91	C	32	β2	-	-
121	C	6	α4	-	Yes
125	C	5	α4	-	-
132	C	53	α4	-	-
161	C	29	α5	-	-
164	C	44	β5	-	-
182	C	31	α6	-	-
190	C	2	β6	-	-
196	C	20	Random coil	-	-
209	C	39	α7	-	-
224	C	34	Random coil	-	-
266	C	17	α9	PLP-Heme cross-talk	-
276	C	38	β9	-	-
317	C	25	α11	-	-
336	C	29	α12	-	Yes
369	C	21	β11	-	Yes
379	C	6	α14	-	-
389	C	62	α15	-	-
413	R	59	Random coil	-	-
439	R	76	α18	-	-
491	R	67	β14	-	Yes
498	R	49	α21	-	-
527	R	-	-	-	-
548	R	76	α22	-	-

[1](N) N-terminal; (C) Catalytic; (R) Regulatory. [2](ASA) Accessible surface area. [3](SS) Secondary structural elements.

Interestingly, among the pathogenic R>C substitutions described in CBS deficient patients, the p.R336C variant is highly frequent in Qatar (Bashir et al., 2015). The finding that all cases belong to three tribes of the Qatari population clearly points towards a founder effect of a *CBS* gene mutation, in combination with a high incidence of consanguineous marriage. In fact, in a study carried out to delineate the genetic and epidemiological aspects of homocystinuria in the Qatari population (El-Said et al., 2006), homozygosity for the mutation c.1006C>T in exon 11 of the *CBS* gene, which encodes for the p.R336C variant, was found in all but one patient investigated, representing 30 out of 31 independent nuclear families. All these 30 families belonged to two single tribes, namely M and K (El-Said et al., 2006).

The c.1006C>T mutation, originating the p.R336C variant, was firstly identified in a patient of English descent (De Franchis et al., 1999) being subsequently observed in patients from Australia (Gaustadnes et al., 2002), Iberian Peninsula (including one homozygous Portuguese patient) (Urreizti et al., 2003) and more recently in a Korean patient (Lee et al., 2005). It is assumed that, most likely, this mutation has recurred independently in different populations and therefore, no clear link could be established between this mutation in the Qatari population and in the other countries, clearly indicating a strong founder effect, making the Qatari population a good study model for the p.R336C variant.

Clinical observations and enzyme measurements in patients with this variant, as well as *in vitro* expression studies, showed that this mutation has a strong impact on enzyme activity and is not associated with vitamin B6-responsiveness (Bashir et al., 2015; Mendes et al., 2015).

The R>C substitution in position 336 of the CBS polypeptide will probably affect the stability of the protein by modifying the hydrogen-bonds established by the different amino acids, as depicted in figure 6A and 6B, thereby modifying also CBS properties, like solubility and activity (Kožich et al., 2010).

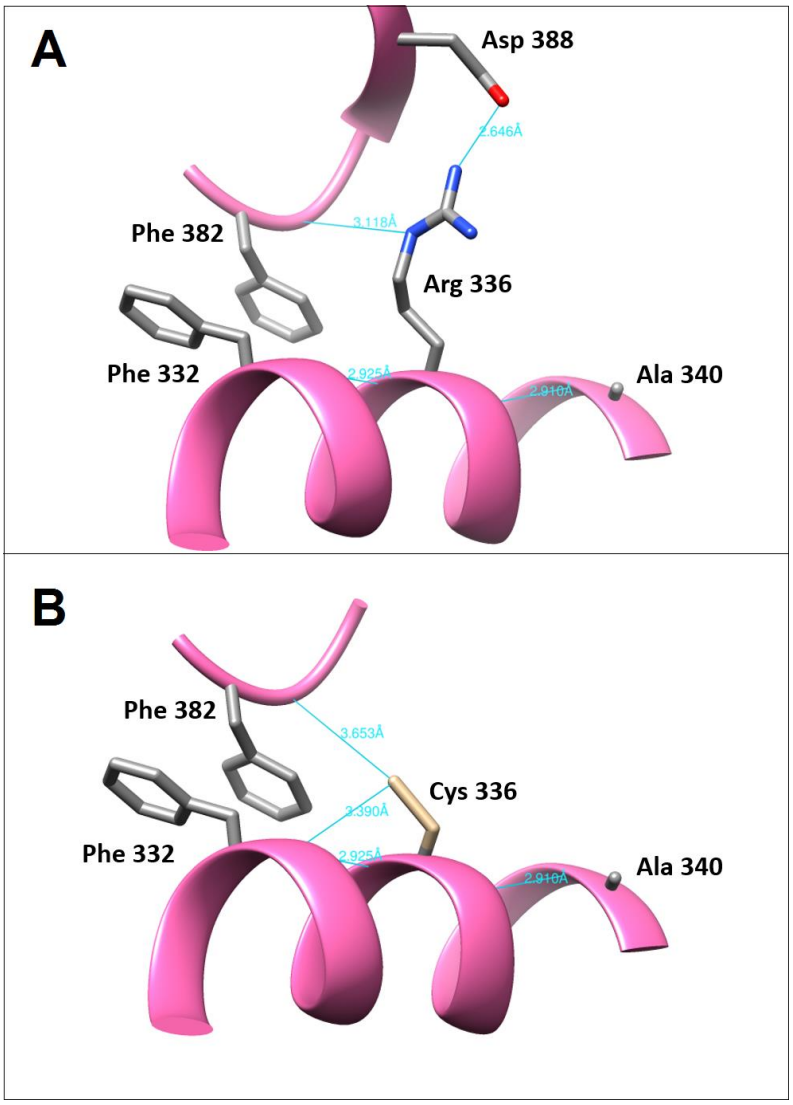

Figure 6 – Representation of the WT Arg336 **(A)** and variant Cys336 **(B)** residues; the H-bonds established by Arg336 and Cys336 with neighboring residues are shown. The images were obtained using the crystal structure of a truncated form of CBS (PDB ID 4L27) and the UCSF Chimera, version 1.8.1 (build 39231)).

1.8. Classical homocystinuria and therapeutic options

At present, the CBS deficiency is usually treated by administration of large doses of vitamin B6 (the PLP precursor) combined with methionine restriction, cysteine supplementation, and enhancement of homocysteine remethylation by betaine and folates (Kožich et al., 2011). As referred previously, vitamin B6 supplementation may be sufficient to treat patients designated as B6-responsive. These patients show a very successful

improvement in their clinical phenotype when receiving pharmacological doses of B6. It is, therefore, recommended that new patients should always be tested for vitamin B6 response. As about 50% of CBS deficient patients are nonresponsive to vitamin B6 and the management of their diet is quite difficult, especially in late-diagnosed cases, new therapeutic options are highly desirable.

Misassembly and aggregation of CBS variants have been described to contribute substantially to the pathogenesis of CBS deficiency. In recent years, chemical and pharmacological chaperones have been reported as a therapeutic option to prevent misfolding or aberrant trafficking of proteins involved in human conformational diseases (Perlmutter, 2002), and as such, these compounds have also been investigated for their ability to facilitate the proper assembly of several CBS variants (Kopecká et al., 2011). The finding that misfolding and decreased activities of some variant proteins may be corrected with the help of chemical chaperones further supports the role of misfolding in the pathogenesis of CBS deficiency and highlight the relevance of identifying pharmacological chaperones (Singh et al., 2007). The proof-of-concept has been already provided, since heme arginate was demonstrated to substantially increase the content and rescue the catalytic activity of seven CBS variants, namely the p.A114V, p.K102N, p.R125Q, p.R266K and p.R369C (Melenovská et al., 2015).

1.8.1. The rescue of p.R336C variant by thiol compounds

It has been shown for two unrelated proteins (ApoE and Factor VIII light chain) that the deleterious effect of R>C substitutions could be functionally overcome by the presence of cysteamine, an aminothiol compound (Horie et al., 1992; Aly et al., 1992). Recently, Mendes and colleagues aimed to study the effect of not only cysteamine (CySH) but also mercaptoethylguanidine (MEG), another thiol compound, in the specific rescue of the p.R336C variant of CBS (Mendes et al., 2015). Cysteamine is the simplest stable aminothiol and a degradation product of Cys (Reid, 1958), while MEG is formed by transguanylation when *S*-(2-aminoethyl)-thiuronium bromide hydrobromide is dissolved in aqueous buffered solutions at neutral pH (Schwartz et al., 1960). Cysteamine has shown not only a neuroprotective effect, but also the capacity to protect the liver against acetaminophen

poisoning via the enhancement of the antioxidant glutathione system (Peterson et al., 1992; Prescott et al., 1976; Di Leandro et al., 2008; Lesort et al., 2003; Pillai et al., 2008). MEG is a selective inhibitor of inducible nitric oxide synthase (iNOS) and a peroxynitrite scavenger (Lohinai et al., 1998). It presents anti-inflammatory properties, including direct inhibition of cyclooxygenase (COX) activity, being also a potent scavenging of hydroxyl radicals (Lohinai et al., 1998).

Mendes and colleagues postulated that these thiols would bind to the mutant residue Cys, forming a structure resembling the wild-type (WT) residue Arg, thereby restoring enzyme activity (Mendes et al., 2013). Interestingly, in all tested conditions the p.R336C variant showed an increase in activity. A 6- and 17-fold increase was observed when samples were incubated for 7h, at 37°C, with 0.25 mM CySH and 1 mM MEG, respectively. No effect was observed for the p.R336H or WT proteins or when Cys was used instead of CySH or MEG. Thermal inactivation profiles were also performed and similar results were obtained: no changes for the WT, but a clear shift of the $T_{1/2}$ of p.R336C to values similar to the WT in the presence of both CySH and MEG. Taking all together, these data support the hypothesis that these compounds specifically rescue the R>C substitution on position 336 of the CBS protein (Mendes et al., 2013).

2. Aims

Presently, a high number of genetic diseases, and in particular Inherited Metabolic Disorders, are classified as Loss of Function Conformational Diseases and are therefore excellent targets for the treatment with compounds able to specifically rescue the misfolded protein (e.g. pharmacological chaperones), increasing their steady-state levels and concomitantly restoring protein function. The features of some of the disease-causing variants found in CBS deficient patients allows also to classify CBS deficiency as a Conformational Disease and CBS variants as potential targets for rescuing with small molecular weight compounds.

Previous work developed by the Metabolism and Genetics group pointed out that the rescue of the p.R336C CBS variant could be achieved by thiol compounds such as CySH and MEG.

Facing the fact that the CBS polypeptide sequence presents 28 Arg residues with five of them being involved in R>C disease-causing substitutions, namely the p.R18C, p.R121C, p.R336C, p.R369C and p.R491C, the main aims of this project were to:

1. Clarify the molecular mechanism underlying the increase in the enzyme activity of several R>C CBS variants, namely R18C, R121C, R164C, R336C, R369C and R491C, in the presence of thiol compounds;
2. Understand whether the localization of the R>C substitution in the protein sequence modulates the observed effect.

3. Material and Methods

3.1. Production of mutant CBS expression constructs

Site directed mutagenesis is a powerful tool to introduce specific DNA mutations. This *in vitro* procedure requires the synthesis of a short DNA primer, containing the desired mutation and a sequence complementary to the template DNA (which must be dam methylated), so it can hybridize with the target DNA sequence of interest.

Recombinant wild-type (WT) and variant CBS proteins were produced using the pET28b expression vector (Clontech Laboratories). The CBS WT cDNA (GenBank IDBC007257.1), cloned into the *Eco*RI and *Xho*I sites of the pOTB7 vector (Thermo Scientific) was used to produce the expression construct (pET28b-6xHis-pepT-hCBSWT) as previously described (Mendes, Colaço et al., 2014). The desired mutations were introduced by site-directed mutagenesis using the NzyMutagenesis Kit (Nzytech) and the specific mutagenic oligonucleotides (Metabion) described in Table 2.

Table 2 – Oligonucleotides used for the site directed mutagenesis reaction

CBS variant	DNA change	Primer	Sequence (5'-3')
p.R18C	c.52C>T	R>C18F	GGCTGCCCCCACTGCTCAGGGCCACACTCG
		R>C18R	CGAGTGTGGCCCTGAGCAGTGGGGGCAGCC
p.R121C	c.361C>T	R>C121F	GGAGCGTGAAGGACTGCATCAGCCGC
		R>C121R	GCGGCTGATGCAGTCCTTCACGCTCC
p.R164C	c.490C>T	R>C164F	GGCAGTGAGGGGCTATTGCTGCATCATCG
		R>C164R	CGATGATGCAGCAATAGCCCCTCACTGCC
p.R336C	c.1006C>T	R>C336F	GGCGTTCACCTTTGCCTGCATGCTGATCGCG
		R>C336R	CGCGATCAGCATGCAGGCAAAGGTGAACGCC
p.L338P	c.9858T>C	L338P-F	CCTTTGCCCGCATGCCGATCGCGCAAGAGG
		L338P-R	CCTCTTGCGCGATCGGCATGCGGGCAAAGG
p.R369C	c.1105C>T	R>C369F	GCAGGAGGGCCAGTGCTGCGTGGTCATTCTGC
		R>C369R	GCAGAATGACCACGCAGCACTGGCCCTCCTGC
p.R491C	c.1471C>T	R>C491F	GCAGTTCAAACAGATCTGCCTCACGGACACG
		R>C491R	CGTGTCCGTGAGGCAGATCTGTTTGAACTGC
p.R491H	c.1472G>A	R>H491F	GCAGTTCAAACAGATCCACCTCACGGACACG
		R>H491R	CGTGTCCGTGAGGTGGATCTGTTTGAACTGC

Note: Mutagenic oligonucleotides are underlined.

The mutagenesis reaction was performed using the buffer supplied by the manufacturer in a final volume of 50 µl containing 125 ng of each primer, 40 µM each dNTPs, 60 ng of template DNA and 2.5 U of NzyProof DNA polymerase (Nzytech). DNA amplification was performed on a PCR Express Thermocycler (Hybaid) using the template program described in Table 3. Whenever necessary, in step 2 the hybridization temperature was changed according to the T_m of the primers (62ºC) as well as the elongation temperature (72ºC).

Table 3 – Amplification program utilized for site directed mutagenesis

Step	Number of cycles	Temperature (ºC)	Time (min)
1	1	95	2
2	18	95	1
		60 (62)[1]	1
		68 (72)[1]	1.5/kb DNA template
3	1	68	15

[1]Changes in temperature.

The DNA was digested with *Dpn*I, to degrade the parental methylated DNA (WT), and the amplified product was introduced into ultra-competent *E. coli* cells (NzyStar cells; Nzytech). The transformed cells were grown on solid Luria-Bertani broth (LB) supplemented with kanamycin (50 mg/L; LB/Kan). After incubation overnight (ON), at 37ºC, one isolated colony was chosen to inoculate 10 mL of LB/Kan which was further incubated ON, at 37ºC, under constant stirring of 140 rpm. The bacterial culture was used to obtain a cellular pellet from which the plasmid DNA was purified.

The authenticity of mutagenesis was verified by DNA sequencing. This technique also allowed excluding for unwanted and unspecific mutagenic events.

3.2. Expression of WT and Variant Proteins in *E. coli*

To obtain the CBS proteins under study in high yields, a prokaryotic expression system and the pET28b expression vector were used. The pET28b vector (Figure 7) presents upstream the multiple cloning site (MCS): (1) the promoter T7 recognized by the RNA polymerase of *Escherichia Coli* phage 7; (2) a six CAT codon sequence, which encodes a small peptide chain of 6 histidine residues (6xHis Tag); (3) a thrombin cleavage site (Leu-Val-Pro Arg-↓-Gly-Ser; pepT) located after the 6xHis Tag and before the recombinant protein starting codon, allowing 6xHis Tag removal; (4) a ribosome binding site (RBS) sequence and (5) the start ATG codon. This vector is further comprised by the T7 terminator (downstream the cloned CBS WT cDNA), the *LacI* gene sequence that assures an efficient promoter repression by the repressor lac, the pBR322 replication origin f1, and the kanamycin resistance gene that assures the selection of this vector using the appropriated culture medium.

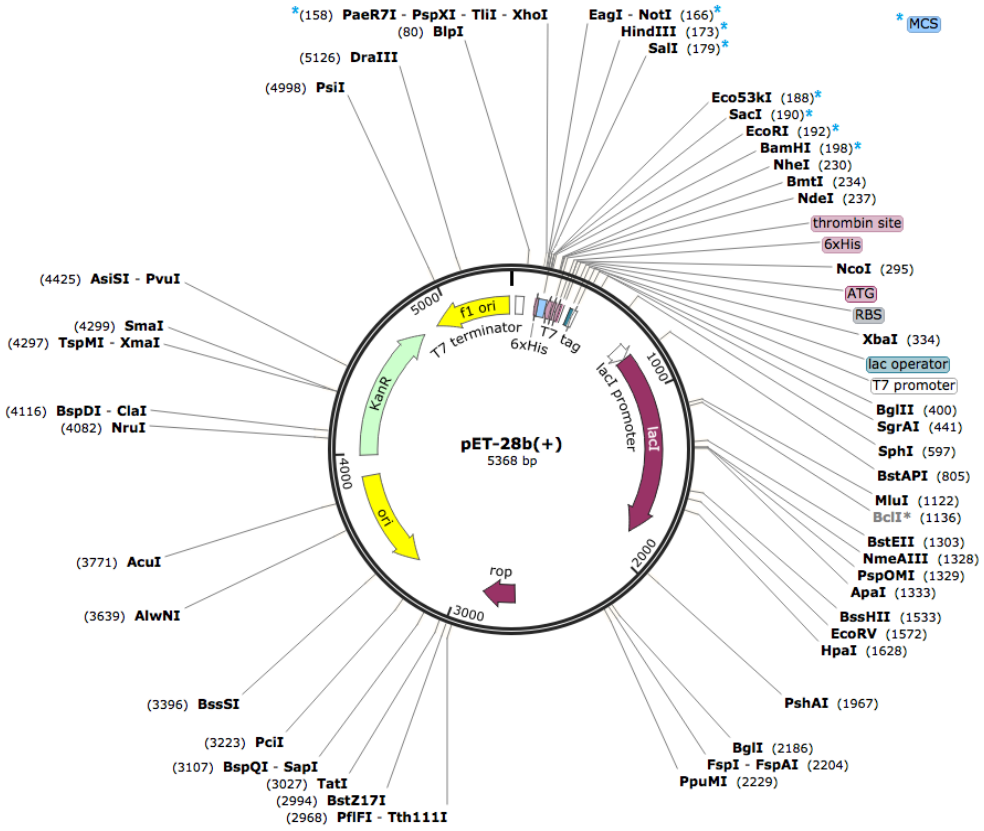

Figure 7 – Schematic representation of the pET28b vector. The figure shows the: Multiple Cloning Site (MCS), T7 promoter, six CAT codon sequence, thrombin cleavage site, Ribosome Binding Sequence (RBS) and the start ATG codon (adapted from (http://www.snapgene.com/resources/plasmid__files/pet_and_duet_vectors_(novagen)/pET-28b(+)/, 14-11-2015)).

The obtained mutant CBS constructs were used to transform competent *E. coli* BL21 (DE3) cells.

To produce the recombinant CBS proteins (WT and variants) 10 mL of LB/Kan was inoculated with an isolated colony. After ON growth, at 37ºC, under constant stirring of 140 rpm, the culture was used to inoculate 1 L of LB/Kan. A supplement of 75 mg of δ-aminolevulinic acid (Sigma-Aldrich) was also introduced in the medium, as this compound is the biosynthetic precursor of Heme (Kang et al., 2012) and essential cofactor for CBS activity and folding (Majtan et al, 2008). The culture was incubated at 37ºC, under a constant stirring of 140 rpm until an Abs_{600nm} of 0.45-0.55. Protein expression was then induced by addition of 0.5 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG; Nzytech). After 4h of incubation at 27ºC under the constant stirring of 140 rpm, the *E. coli* cells were harvested by centrifugation at 2,057 g, for 10 min, at 4ºC. The obtained cellular pellet was stored at -20ºC, until usage (no longer than 1 day).

3.3. Purification of WT and Variant Proteins

The selective isolation and purification of enzymes and other biologically important macromolecules by affinity chromatography exploits the unique biological property of these proteins to bind ligands specifically and reversibly (Cubtrecasas et al., 1968). The protein to be purified is passed through a column containing an insoluble polymer or gel to which a specific competitive inhibitor or ligand has been covalently attached. Proteins not exhibiting appreciable affinity for the ligand will pass through the column, whereas those which recognize the inhibitor will be retarded to an extent related to the affinity constant under the experimental conditions. In this work, the recombinant proteins present N-terminally a 6xHis tag with high affinity for metal ions (such as Ni^{2+}), allowing protein purification by immobilized metal affinity chromatography (IMAC) using a Ni^{2+}-chelating resin (Ni-NTA; Qiagen). To purify the expressed proteins, the *E. coli* BL21 (DE3) cells were resuspended in 10 mL (for 1 L of culture) of Buffer A (50 mM potassium phosphate, 300 mM KCl, pH 7.0), 10% glycerol, 1 mg/mL lysozyme (AppliChem), 1 mM phenylmethylsulfonyl fluoride (PMSF) (Merck) and a small amount of DNAse (AppliChem). After 30 min incubation on

ice, the cells were disrupted by 3 cycles of ultra-sonication during 60 seconds, with a 50% duty free cycle (Vibra Cell sonicator; Sonics Materials). The cell lysate was centrifuged at 12,857 g, for 40 min, at 4ºC, obtaining an insoluble (pellet) and a soluble (supernatant) fraction. To the soluble fraction, imidazole (Merck) and β-mercaptoethanol (Applichem) were added to a final concentration of 10 mM each. The Ni-NTA resin (500 µL resin per 10 mL of soluble lysate fraction) was previously equilibrated with successive washes of: 1 volume Milli-Q water (3x) and 1 volume Buffer A containing 10 mM imidazole (2x). Between washes, the resin was centrifuged at 425 g, for 2 min, rejecting the supernatant. After resin equilibration, the soluble lysate fraction was added to the equilibrated resin and the suspension was incubated at 4ºC, for 1 h, under constant stirring of approximately 100 rpm. After this time period, the solution was loaded into a polypropylene conic column (poly-prep; Bio-Rad) and the flow-through was collected. The purification process was performed at 4ºC, using a gradient of Buffer A with increasing imidazole concentrations: 20 mM (2x 5 mL), 50 mM (2x 5 mL), 75 mM (5x 500 µL) and 500 mM (5x 500 µL). To remove the imidazole from the samples, the fractions corresponding to the eluted protein (500 mM) were pooled and loaded into a PD-10 Column (GE-Healthcare) previously equilibrated with Buffer A. The recombinant proteins were recovered with Buffer A and quantified using the Bradford colorimetric reagent (Bradford MM, 1976) from Bio-Rad and Bovine serum albumin (BSA; Calbiochem) as the standard at the final concentrations of 1, 2, 4, 8, and 10 µg/mL. Absorbance was measured at 595 nm in a Hitachi spectrophotometer (Model U-2000). All assays were performed in triplicate to assure precision.

When necessary samples were concentrated by ultrafiltration using an Amicon Ultra 15 centrifugal filter device (MWCO 30 kDa; Millipore). The samples were centrifuged at 3,220 g, for 5 min, at 4ºC, until a final volume of approximately 1 mL. Between each centrifugation, the solution was softly resuspended to desorb protein that might accumulate at the filter membrane. The obtained protein was stored at -80ºC until its usage.

3.4. Characterization of recombinant proteins

3.4.1. SDS-PAGE Analysis

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) is a technique performed on denaturing polyacrylamide gels. In this system, the proteins are separated according to their molecular mass. SDS-PAGE is a powerful tool to evaluate not only the protein's molecular mass, but also the purity grade of protein extracts. SDS is a negatively charged detergent that dissolves hydrophobic molecules. Thus, SDS denatures the proteins and confers them a negative charge allowing the proteins to move from the cathode (negatively charged) to the anode (positively charged).

In our work, the discontinuous system of Laemmli (Laemmli, 1970) was used. This system utilizes a stacking gel (4%) and a resolving gel (different concentrations) and three different buffers: (i) a stacking buffer (Tris.HCl 0.125 mM, pH 6.8, SDS 0.1%); (ii) a resolving buffer (Tris.HCl 0.375 mM, pH 8.8, SDS 0.1 %) and; (iii) a running buffer (Tris 3 g/L; glycine 14.4 g/l; SDS 0.1 %, pH 8.3). This discontinuous Tris-glycine buffer system is the most suitable for separation of proteins with molecular masses exceeding 20 kDa. Sample separation was performed on a mini-vertical electrophoresis system Hoefer SE245 (GE Healthcare).

The samples were applied in a final volume of 20-25 µL, being previously prepared in loading buffer (Tris.HCl 0.24 M pH 6.8, glycerol 40%, SDS 8%, bromophenol blue 0.02%, β-mercaptoethanol 25%) and denatured at 95ºC for 5-10 minutes.

After electrophoresis, the proteins were visualized by staining with Brilliant Blue Coomassie R250 (Sigma). To evaluate the molecular mass of the recombinant proteins, protein molecular mass markers were used, namely the Low Molecular Weight (LMW) Protein Marker (Nzytech; 18.5, 26, 32, 40, 48, 66 and 96 kDa) and NzyColour Protein Marker II Prestained Protein Ladder (Nyztech; 11, 17, 20, 25, 35, 48, 63, 75, 100, 135, 180, and 245 kDa).

The purity grade and molecular mass of the produced recombinant proteins were evaluated using a 10% polyacrylamide resolving gel. Protein purity was assessed by densitometric analysis, using the ImageJ 1.42q software (National Institutes of health; USA).

3.4.2. Assay of Enzymatic Activity

The enzyme activity of recombinant CBS was determined in a reaction volume of 110 µL, containing 100 mM Tris HCl pH 8.6, 5 mM L-Ser and 5 µg of WT CBS or 10 µg of the variant forms. To prevent the unspecific formation of Cth, 5 mM EDTA (Smith et al., 2012) were also included in the reaction mixture. The reaction was started by the addition of 2.7 mM L-Hcy, prepared from L-Hcy thiolactone hydrochloride (Sigma) (Mendes, Colaço et al., 2014). After 30 min at 37°C, the enzymatic reaction was stopped by the addition of 110 µL ethanol:acetic acid (98:2) solution. After centrifugation at 10,621 g, during 5 min, at 4ºC, the supernatant was used to determine the amount of the produced L-Cth. The enzymatic activities were expressed as nmol of Cth formed per mg of protein per hour at 37ºC (nmol $Cth.h^{-1}.mg^{-1}$). The enzymatic assays were performed in the absence or in the presence of 1 mM PLP or 1 mM MEG (pre-incubation at 37ºC for 1h). Three independent experiments were performed for all the assays.

3.4.3. Quantification of Cystathionine by HPLC

To determine the activity of the recombinant purified CBS proteins, it was necessary to develop a method to quantify the L-Cth produced during the enzyme reaction. High pressure liquid chromatography (HPLC) with fluorimetric detection is usually associated with high sensitivity, linearity and reproducibility. Therefore, due to the availability of this technique in our laboratory (Met&Gen group), a method to detect L-Cth by HPLC after derivatization with *O*-phthaldialdehyde (OPA) and fluorimetric detection was implemented.

The amount of L-Cth produced after 30 min was quantified as previously described (Ravanel et al., 1995), using an HPLC Waters System equipped with a fluorimeter, and a Gemini C18 column (150x4.6 mm, 5 µm particle size; Phenomenex). The mobile phase consisted of buffer B (85 mM sodium acetate, pH 4.5; 6% acetonitrile) and Buffer C (60% acetonitrile) pumped at 1 $mL.min^{-1}$ flow rate. A linear gradient of 40 to 80% solvent C (0-6 min), 80% solvent C (6-9 min), 80-40% solvent C (9-10 min), 40% solvent C (10-12 min) was used. The injected volume was 5 µL and the OPA derivatives were detected by measuring the fluorescence at λ_{exc}= 340 nm and λ_{em}= 455 nm. During running, the column was maintained at 40ºC.

Cystathionine solutions at different concentration (7.81, 15.6, 31.25, 62.5, 125, 250 and 500 µM) were used to determine the detection limit and linearity of the method. Norvaline was used as the internal standard.

3.4.4. Limited Proteolysis

Proteolysis of native proteins can be limited, and the rate of digestion of peptide bonds depend on several parameters, the most important of which is the capability for local flexibility of the polypeptide chain in the region of the scissile bond (Hubbard, 1998). Proteolytic enzymes can therefore be used as probes of flexible or exposed regions of a polypeptide chain, including conformational shifts as those elicited by ligand binding (Fontana et al, 2004; Ellison et al, 1995). Usually proteases, such as trypsin, with a wide proteolytic activity, should be used.

Limited proteolysis of WT and CBS variants was performed at 25ºC in buffer A, with a final protein concentration of 0.6 mg/mL and a trypsin (Sigma):CBS ratio of 1:100 (mass/mass), in the absence or presence of 1 mM cysteamine, 1 mM MEG or 2.7 mM PLP. At different time points (0, 2, 5, 10, 15, 30, 45, 60 and 120 min), 20 µL aliquots of the reaction were mixed with soybean trypsin inhibitor (Sigma) at a protease:inhibitor ratio of 1:1.5 (by mass). Samples were loaded on a 10% polyacrylamide gel as described above. Gels were stained with Coomassie brilliant blue R250 and the band corresponding to the full-length protein was quantified by densitometry using ImageJ 1.42q software.

3.4.5. Differentiation Scanning Fluorimetry

Differential Scanning Fluorimetry (DSF) is a thermal shift assay where the temperature at which a protein goes from well-defined structure to a disorder state is determined. This technique is based on monitoring changes in the fluorescence signal of SYPRO orange dye, as it interacts with the hydrophobic residues of a protein undergoing thermal unfolding (Figure 8) and allows determination of the melting temperature (T_m), the midpoint of the protein melting curve. The dye's fluorescence signal is quenched in the aqueous environment of a properly folded protein, but becomes unquenched when exposed

to the protein's hydrophobic core upon unfolding. The T_m will increase in the presence of ligands that bind to the protein stabilizing it.

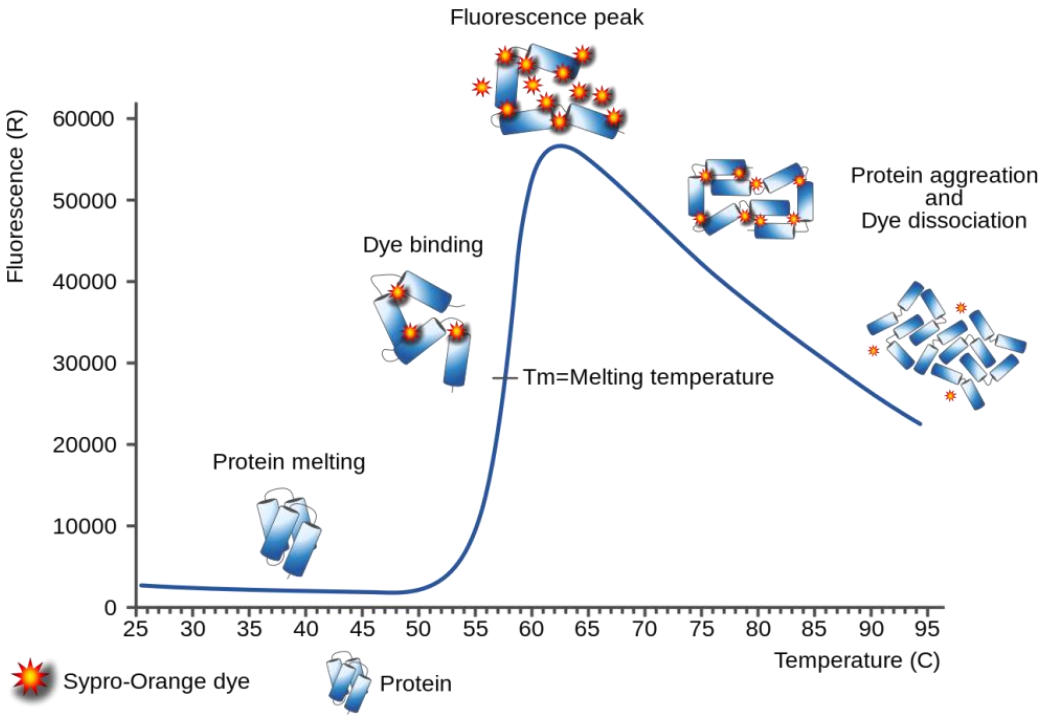

Figure 8 – Schematic representation of the Thermal Shift Assay (adapted from Argonne National Laboratory, Biosciences Division - Molecular and Systems Biology; http://www.bio.anl.gov/molecular_and_systems _biology/Sensor/sensor2.html, 06-08-2015).

The WT and CBS variants were centrifuged (10,621 g, for 10 min, at 4ºC) to remove insoluble aggregates. Proteins were diluted in buffer A to a final concentration of 0.05 µg/mL. Sypro Orange (Invitrogen; 5000× stock solution) was used at a final 2.5× working concentration. Stabilizing compounds (1 mM Cysteamine, 1 mM MEG and 1 mM PLP) were tested without and with pre-incubation with the proteins, in Buffer A, for 1h at 37ºC.

The DSF assay was performed on a C1000 Touch thermal cycler equipped with a CFX96 optical reaction module (Bio-Rad) in a final volume of 50 µL. The PCR plate was sealed with Optical-Quality Sealing Tape (Bio-Rad) and centrifuged at 500xg, at 4 ºC, for 3 min. DSF was carried out by increasing the temperature from 20 to 90 ºC, at 1 ºC/min. The obtained results were analyzed using GraphPad Prism 6 software (GraphPad, Software, Inc).

3.4.6. Modification of Recombinant proteins with PEG-Mal

Polyethylene glycol (PEG) is a molecule frequently used as a protein PEGylation reagent. PEGylation is the process of covalently attach the polyethylene glycol polymeric chain – which can have different molecular weight – to another molecule, such as proteins. PEGylation was originally designed to improve the pharmacokinetic profile of protein therapeutics (Fried et al, 2002; Hadziyannis et al, 2004; Talal et al, 2006). In our work, PEG-Maleimide (PEG-Mal) was used to covalently attach PEG to the Cys residues of the CBS proteins through the maleimide moiety. In fact, N-ethylmaleimide (NEM) is an alkylating reagent that reacts with sulfhydryls (–SH; such as that found in Cys) to form stable thioether bonds (Smyth D et al., 1960).

In our work, PEG-Mal with a polymeric chain of 5000 (PEG_{5000}) was used (Lu et al., 2001). The PEG-Mal-modified protein migrates more slowly on SDS-PAGE than those of the unmodified protein, allowing to monitor for the number of available Cys residues accessible to modification by PEG-Mal.

PEG-Mal (Sigma-Aldrich) was freshly prepared in buffer A. In all cases, samples were subjected to a pulse of a molar excess of PEG-Mal. Different temperatures and incubation times were tested to optimize the reaction.

To test MEG binding to CBS free Cys residues, a pre-incubation with 1 mM MEG for 1h at 37ºC was performed. MEG was further removed from the sample using a MiniTrap G-25 column (GE-Healthcare), and the resulting protein was then PEGylated with PEG-Mal.

Samples were loaded on 7.5% or 10% denaturant polyacrylamide gel as described above. Gels were stained with Coomassie brilliant blue R250, and the band corresponding to the full-length protein was quantified by densitometry using ImageJ 1.42q software.

3.4.7. Evaluation of –SH groups by DTNB

Following the PEG-Mal modification, the free Cys residues were evaluated using the Ellman's reagent, 5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB). This reagent is often used to determine the number of free –SH groups in proteins, cells and plasma by absorption measurements (Ellman, 1959; Robyt et al, 1971; Sedlak et al, 1968). The procedure is based on the reaction of the thiol group with DTNB (Figure 9) to produce an equivalent amount of

the yellow chromophore 5-thio-2-nitrobenzoic acid (TNB^{2-}), which absorbs at 412 nm. However, it must be noticed that the DTNB method could underestimate the concentration of free thiol groups in a protein due to its incomplete reaction.

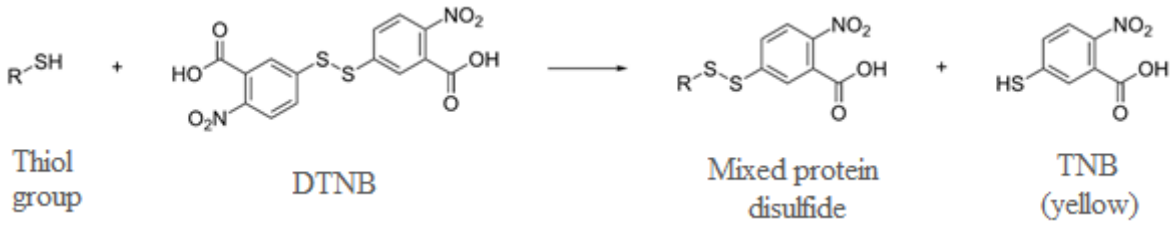

Figure 9 – Reaction of DTNB with a thiol compound (R–SH).

The DTNB (Sigma-Aldrich) solution was prepared at 1 mM in Phosphate buffer 50 mM, pH 8.0. The reaction was performed in a final volume of 1 mL, containing DTNB at 100 μM (final concentration). Samples (before and after PEG-Mal modification) were incubated at RT, for 5 min. The absorbance was measured at 412 nm. Assays were performed in triplicate.

4. Results and Discussion

4.1 Characterization of the R>C CBS variant proteins

As referred previously, CBS contains in its sequence 28 Arg residues, which are localized all over its sequence (Table 1; Figure 3A). In order to understand if the rescue of the activity of R>C CBS variants depends on the localization of the affected residue, six amino acid residues presenting different ASA values and domain localization were selected for mutagenesis, namely the R18, R121, R164, R336, R369 and R491 (Figure 3B). Additionally, an equivalent R>H substitution in position 491 was also studied as a control (p.R491H).

The variant proteins were produced in a prokaryotic expression system, already in use in the laboratory. Functional and structural studies were performed with the purified recombinant proteins in the absence and in the presence of aminothiol compounds, namely CySH and MEG.

4.1.1. Production of the R>C and R>H expression constructs

The desired mutations were successfully introduced in the CBS cDNA cloned in the expression vector pET28b(+) (Invitrogen) by site-directed mutagenesis using the template program described in the Material and Methods section (Table 3). However, for the introduction of the p.R121C and p.R164C mutations an adjustment to the hybridization and elongation temperature was necessary, since no colonies were obtained after *E. coli* transformation. Following DNA purification, sequencing analysis showed the correct introduction of the desired DNA changes, which are presented in Figure 10. The complete sequencing of the mutated cDNAs further allowed verifying that no additional mutations were introduced (data not shown).

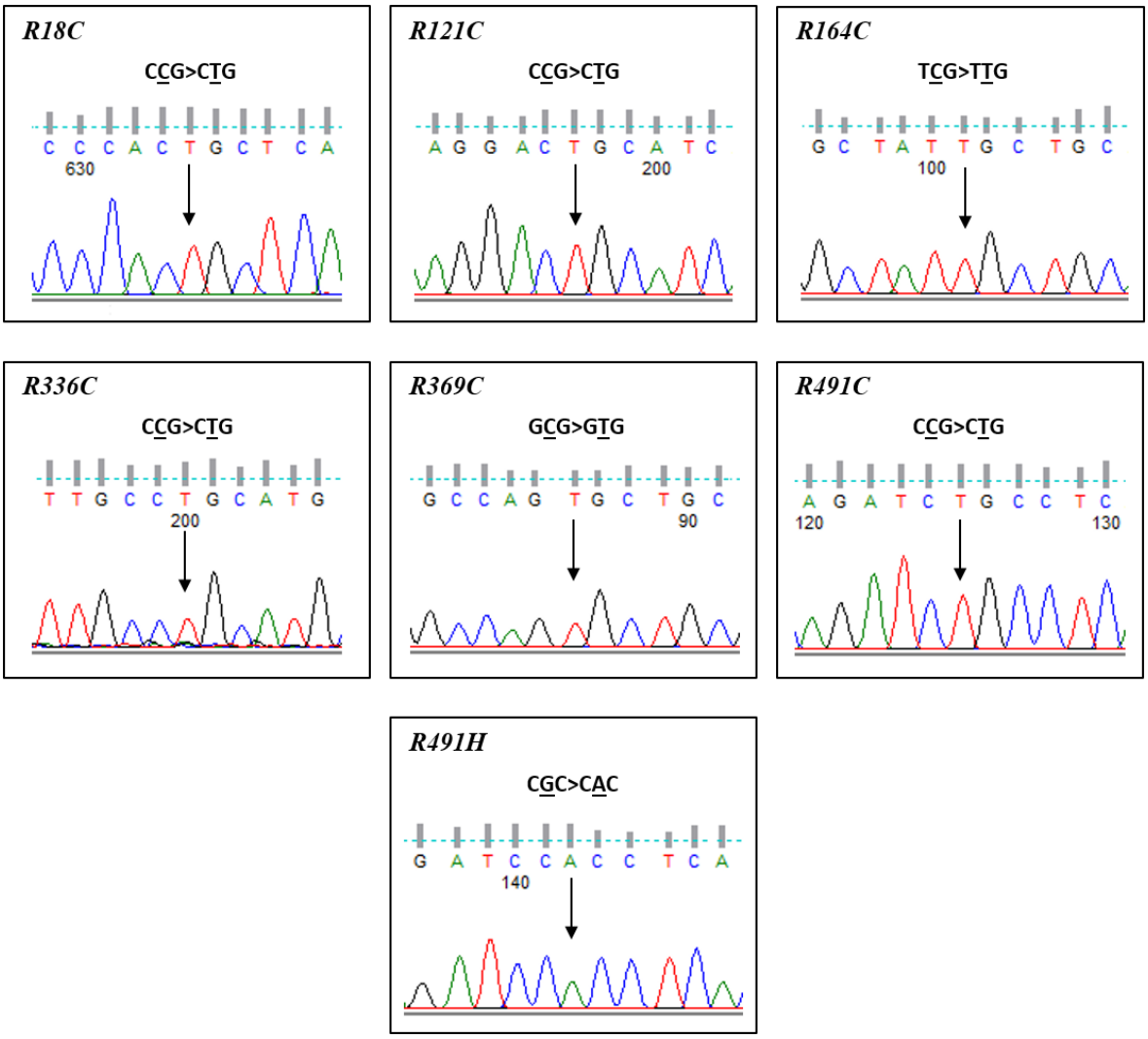

Figure 10 – Partial sequence panels of the CBS cDNA after site-directed mutagenesis. Arrows indicate the position of the single base substitution in the cDNA sequence. For each mutation, the wild-type and mutated codon is shown with the base change underlined.

4.1.2. Expression and purification of recombinant WT and CBS variants

Using the expression conditions described in Material and Methods section, we were able to produce the recombinant CBS variant proteins in the soluble fraction of *E. coli* lysates. In Figure 11, an SDS-PAGE representative of the elution profile of a CBS variant protein (p.R18C) is shown. The CBS proteins were eluted with 500 mM imidazole mostly in the first three fractions (Figure 11, lanes 4-6).

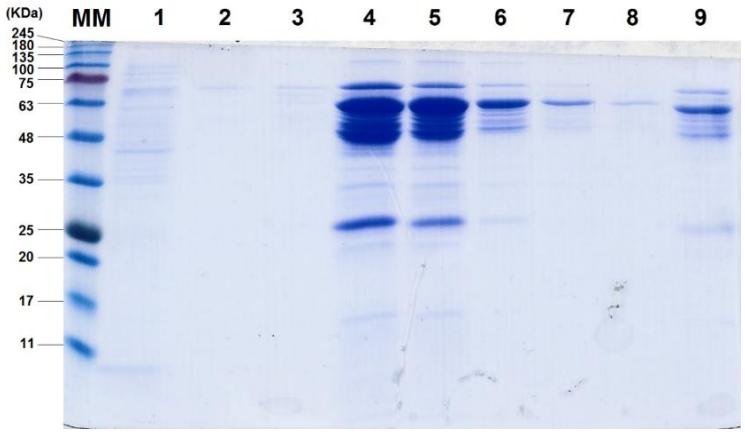

Figure 11 – SDS-PAGE analysis of the purification of p.R18C variant. The recombinant CBS proteins were purified by immobilized metal affinity chromatography using a gradient of imidazole concentration: 10 mL of 20 mM (1), 10 mL of 50 mM (2), 2.5 mL of 75 mM (3), and 5 x 500 µL of 500 mM (4-8). (1) NzyColour protein marker II (Nzytech); (9) Purified CBS recombinant protein after removal of imidazole by gel filtration (using a PD-10 column). In each lane an aliquot of 16 µL of each imidazole elution fraction was applied.

For all the expressed CBS proteins, the molecular mass of the main band corresponded to 63 kDa (Figure 12), which is the estimated molecular mass of the recombinant CBS proteins. Interestingly, for all the recombinant CBS proteins (WT and variants), after gel filtration chromatography we could observe a decrease of the contaminating proteins resulting in a high purity grade (Figure 11, lane 9).

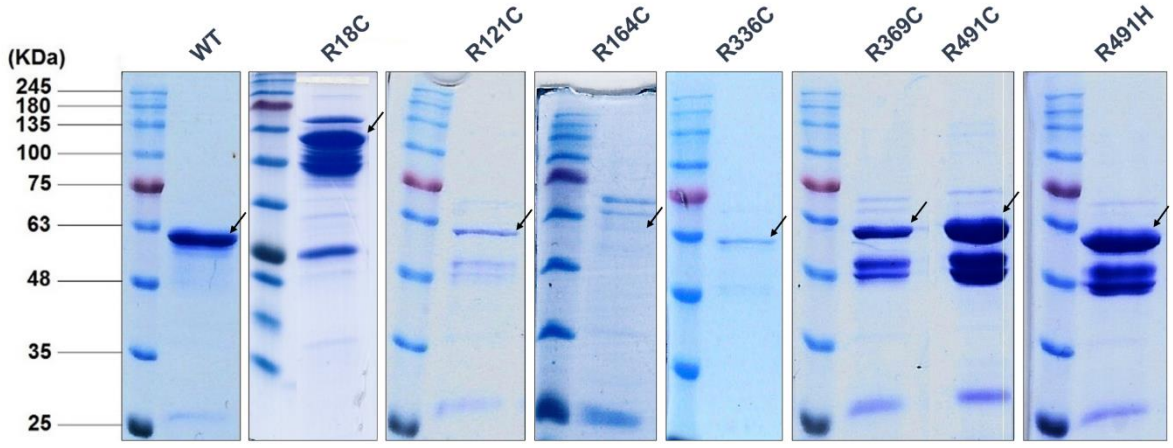

Figure 12 – SDS-PAGE analysis of the purified CBS recombinant proteins (WT and variants) after immobilized metal affinity chromatography (IMAC). The recombinant CBS proteins are indicated by an arrow.

However, when comparing with the WT protein, the purity grades and yields of the CBS variants varied significantly from 7% (p.R164C) to 58% (p.R121C) (Figure 12 and Table 4). Due to the low yield and low purity grade of the p.R164C variant, no further studies were performed for this protein.

Table 4 – Yield and purity grade of the produced recombinant CBS proteins after IMAC purification

	Protein (μg/mL)	Yield (mg/L culture)	Purity grade (%)
WT	597.94	478.35	80
p.R18C	674.76	256.41	38
p.R121C	475.24	275.64	58
p.R164C	303.17	21.22	7
p.R336C	529.24	278.04	48
p.R369C	478.79	220.24	46
p.R491C	662.67	271.69	41
p.R491H	811.71	373.39	46

4.1.3. Conformational flexibility of the R>C recombinant proteins

The susceptibility of WT and CBS variants to limited proteolysis by trypsin was followed over time. This allowed us to probe the changes occurring, not only in the native folding of the enzyme which are caused by the mutations, but also to assess the effect of two aminothiol compounds on the tertiary structure of the variant proteins. As such, the CBS WT and R>C variant proteins were incubated for 1h at 37°C in the presence and absence of CySH and MEG. As a control, the p.R491H was also tested. After incubation, trypsin was added and at different time points, aliquots were collected and analyzed by SDS-PAGE, as shown in Figure 13. The percentage of remaining full-length enzyme was followed by densitometric analysis. The obtained values were plotted and fitted to a single exponential decay equation (Figure 14), allowing to obtain the decay constant of proteolysis (k_P) and the $t_{1/2}$ (time at which 50% of the full-length protein remains intact) (Table 5).

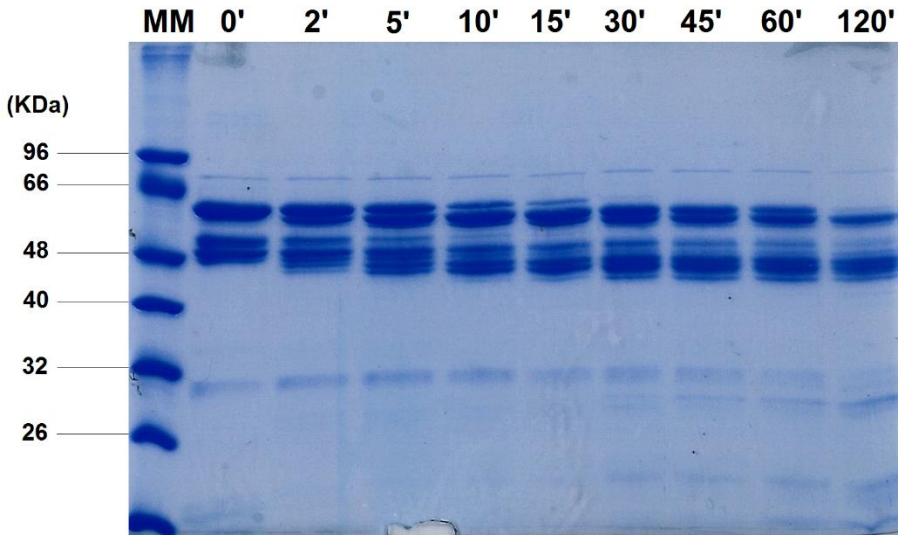

Figure 13 – SDS-PAGE analysis of the time-course for the limited proteolysis by trypsin of p.R491C variant. At different time points (0 to 120 minutes) aliquots were taken and analyzed on a 10% SDS-PAGE. See text for details.

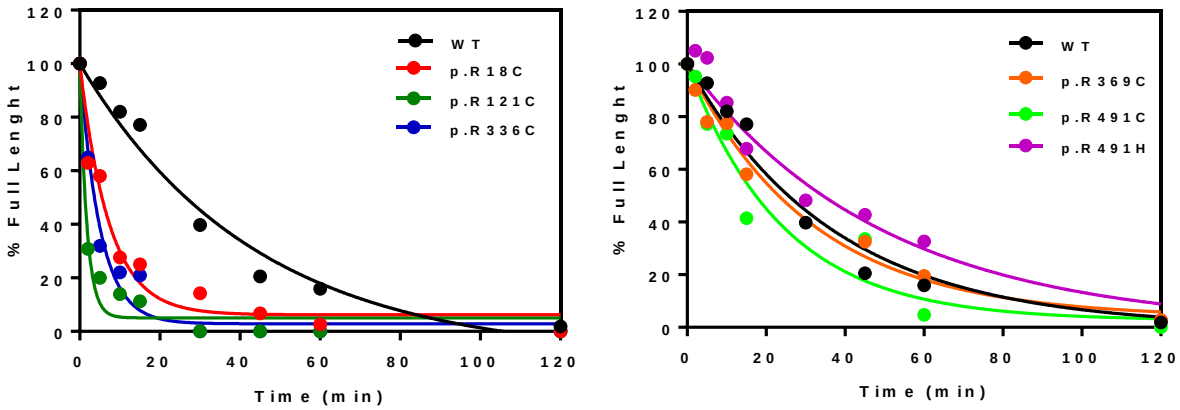

Figure 14 – Time-course for the limited proteolysis by trypsin of WT CBS and variant proteins. Data represent the mean of two independent experiments. Data were fitted to a single exponential decay equation to obtain the decay constant of proteolysis (*k*p). The polyacrylamide gel presents the aliquots from the different time points (0, 2, 5, 10, 15, 30, 45, 60 and 120 min).

In the absence of the studied aminothiol compounds, and when comparing to the WT, the R>C variant proteins p.R18C, p.R121C and p.R336C presented an increase in the k_P, with the corresponding decrease in $t_{1/2}$ (Figure 14, Table 5) value.

Table 5 – Proteolysis rate (*kp*) and half-life ($t_{1/2}$) of WT and variant forms of CBS

	kp (min^{-1})			$t_{1/2}$ (min)		
	-	**CySH**	**MEG**	**-**	**CySH**	**MEG**
WT	0.027	0.39	0.038	25.6	2.70	24.07
R18C	0.14	0.052	0.054	5.00	13.2	12.8
R121C	0.57	0.25	0.46	1.22	2.74	1.49
R336C	0.20	0.24	0.046	3.50	2.84	17.24
R369C	0.032	0.15	0.077	21.9	4.66	9.02
R491C	0.034	0.089	0.026	26.7	7.75	27.1
R491H	0.020	0.071	0.042	34.4	9.74	16.71

Interestingly, the p.R369C and p.R491C variants showed a mild change in the k_P, while the p.R491H presented a lower k_P value (Figure 14, Table 5). As expected, and taking into consideration that the R>C variants selected for our study are disease-causing mutations, these data indicate that a change in the native folding and/or conformational flexibility of the enzymes, caused by the amino acid change, might have occurred.

After incubation with the aminothiol compounds CySH and MEG, the variant proteins showed a change in *kp* and $t_{1/2}$ with different trends for different variants. As shown in Table 5 and Figure 14, in the presence of CySH the WT and the majority of the variant proteins (except the p.R18C and p.R121C) increased their k_p and consequently decreased their $t_{1/2}$ indicating a higher susceptibility for proteolysis in the presence of this aminothiol.

However, in the presence of MEG only the p.R18C and p.R336C decreased their k_p significantly, while the p.R121C showed only a moderate decrease, suggesting that the aminothiol MEG may stabilize these variants. The p.R369C and p.R491H variants showed a lower degree of proteolysis (higher k_p), while the WT and p.R491C did not change significantly (Table 5, Figure 15).

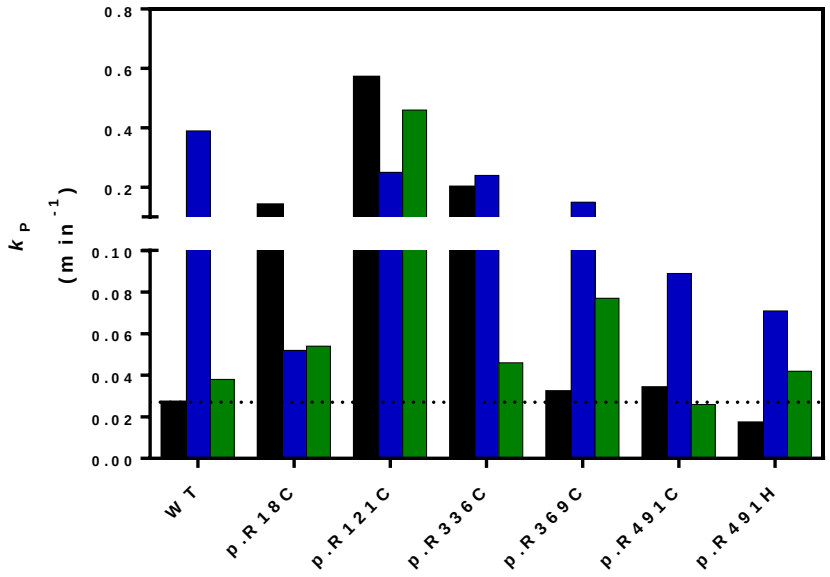

Figure 15 – Proteolysis rate of the WT and variant proteins in the absence (black columns) and in the presence of 1 mM CySH (blue columns) or 1 mM MEG (green columns). The dotted line indicates the k_p for the WT CBS in the absence of thiol compounds. See text for details.

The data gathered indicate that incubation with the studied aminothiol compounds lead to a change in the conformation of the studied variant proteins probably by an interaction of these compounds with specific R>C residues, thereby altering their proteolysis rate, and thus leading to more stabilized or more destabilized proteins, depending on the localization of the affected residue.

4.1.4. Thermal stability of the R>C recombinant proteins

Differential scanning fluorimetry (DSF) was employed to study the thermal stability of the WT and variant CBS proteins. Using DSF and a specific fluorescent dye (Sypro orange), which binds to hydrophobic residues, we were able to follow protein unfolding and determine the respective T_m, i.e. temperature at which half of the protein is unfolded (Figure 16).

The DSF thermal denaturation curves for WT and variant CBS proteins show two transitions (Figure 16), allowing the estimation of T_{m1} and T_{m2} corresponding to the thermal denaturation of the C-terminal and catalytic domain, respectively (Table 6).

In the three assay conditions tested – in the absence of aminothiol compounds, in the presence of aminothiol compounds with pre-incubation, and in the presence of aminothiol compounds but without pre-incubation – all the studied variants presented T_{m1} and T_{m2} values lower than the WT ($|\Delta T_m| > 2^oC$). The only exception was the p.R369C variant that presented a decreased T_{m1} but a normal T_{m2}. From the thermal stability data gathered in the presence of the studied aminothiol compounds for the CBS WT and variants, it is possible to observe that only the p.R336C variant is sensitive to the presence of the aminothiol compound MEG, increasing the stability of the catalytic domain (corresponding to T_{m2}).

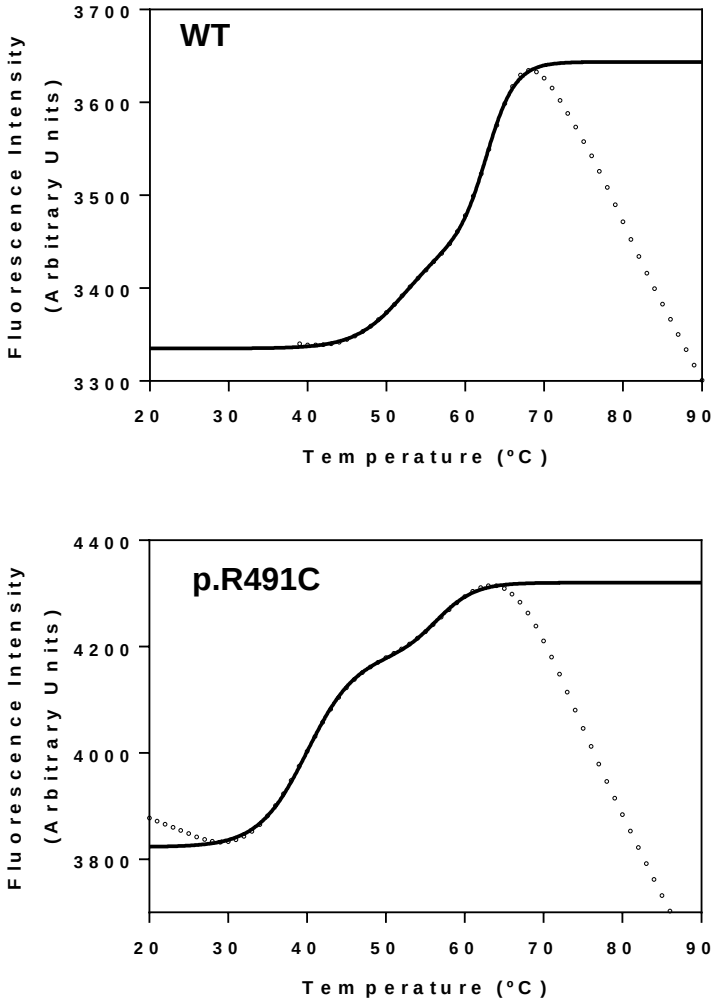

Figure 16 – Thermal unfolding of WT and p.R491C CBS proteins monitored by differential scanning fluorimetry (DSF). Data represent the mean of three assays. Experimental data is represented by the open symbols and the line corresponds to non-linear regression analysis using a biphasic equation. See text for details.

Table 6– Melting temperature (T_m) of WT and variant forms of CBS, in the absence (-) and presence of 1 mM of CySH or MEG, determined by differential scanning fluorimetry

	No pre-incubation						Pre-incubation					
	T_{m1}			T_{m2}			T_{m1}			T_{m2}		
	(-)	CySH	MEG	-	CySH	MEG	(-)	CySH	MEG	-	CySH	MEG
WT	52	52	-	63	64	-	52	48		63	60	-
p.R18C	-	-	-	-	-	-	-	-	-	-	-	-
p.R121C	44	44	44	51	51	51	45	45	44	52	51	52
p.R336C	41	-	43	54	-	60	43	44	44	51	51	60
p.R369C	47	45	45	60	58	58	49	47	49	61	59	59
p.R491C	39	42	42	57	58	59	40	42	42	59	60	59
p.R491H	-	-	-	-	-	-	-	-	-	-	-	-

4.1.5. Assessment of MEG binding by PEG-Mal modifications

As stated in Material and Methods section, the assessment of the protein free –SH groups was performed using PEG-Mal. This polymeric compound is able to react covalently with free –SH groups, leading to PEGylated forms of the protein with an increase of ~5 kDa for each molecule of PEG-Mal added. This technique was used to determine if the tested aminothiol compounds were able to specifically react with the Cys residues of the R>C variant proteins, thus decreasing the available free –SH groups. However, to the limited amount of purified CBS proteins (WT and variants), an optimization of the PEG-Mal modification was firstly performed using instead a recombinant form of the wild-type S-adenosyl-homocysteine hydrolase (SAHH).

4.1.5.1 Optimization of PEG-Mal modification with SAHH

As referred above, to optimize the conditions of the PEG-Mal reaction, and in order to save the recombinant CBS protein for later studies, a recombinant protein available in the laboratory, SAHH, was used as it has a high expression rate. The protein SAHH has 10 Cys residues in its structure (similar to CBS, which presents 11 Cys residues) and a molecular mass of 43 KDa. After SAHH PEGylation, 10 different protein forms with MM ranging from 48 to 93 kDa were expected (48, 53, 58, 63, 68, 73, 78, 83, 88 and 93 kDa).

The first parameter to be optimized was the incubation time. For this, we used a temperature reaction of 4ºC, constant stirring of 140 rpm and 1:400 protein:PEG-Mal ratio (by mass) during different reaction times (0, 30, 60 and 120 minutes). As shown in Figure 16, the SAHH protein immediately reacted with PEG-Mal, leading to additional 5 bands of increasing MM, visible in a 10% SDS-PAGE. The band corresponding to the WT SAHH decreases its intensity with time, while additional bands of higher MM became visible as PEG-Mal molecules bind to a Cys residue. At 120 min, a total of 6 additional bands were observed (Figure 17).

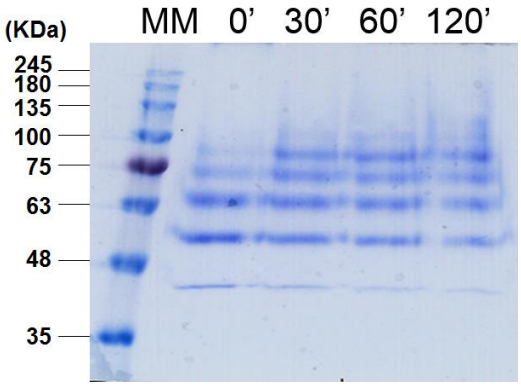

Figure 17 – SDS-PAGE analysis of PEG-Mal binding to S-adenosyl-homocysteine hydrolase (SAHH) at 4ºC. SAHH was incubated with 1:400 protein:PEG-Mal ratio (by mass) at 4ºC for 0, 30, 60 and 120 minutes, under constant stirring of 140 rpm. The unmodified SAHH protein presents a molecular mass of 43 kDa.

Interestingly, it is possible to observe that the binding reaction of PEG-Mal to protein seems to be instantaneous, as observed for the presence of various bands of higher molecular mass (Figure 17, lane 0’ and Figure 18, lane 0’) than the SAHH protein (43 kDa; Figure 18, lane 0’*).

In order to increase protein PEGylation rate, we then tested a higher temperature reaction (25ºC; RT) and more time-points. As shown in Figure 18, at RT a higher number of additional bands was obtained after 10 min of reaction. After this time period, the bands intensity decreased probably due to the aggregation of the protein.

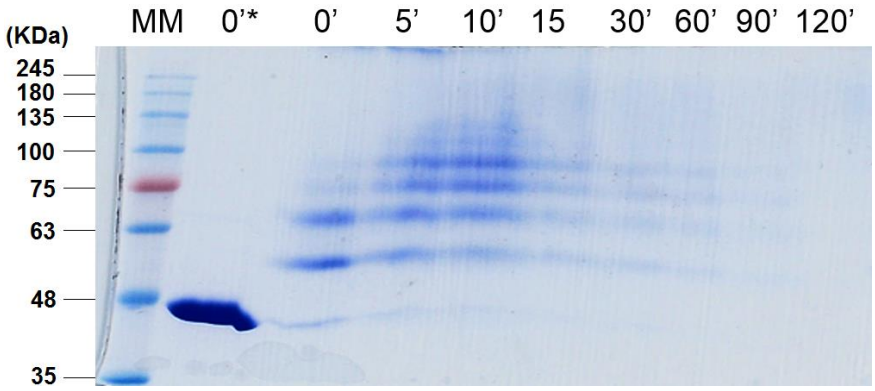

Figure 18 – SDS-PAGE analysis of PEG-Mal binding to S-adenosyl-homocysteine hydrolase (SAHH) at 25ºC (RT). SAHH was incubated with 1:400 protein:PEG-Mal ratio (by mass) at RT for 0, 5, 10, 15, 30, 60, 90 and 120 minutes, under constant stirring of 140 rpm. 0’* represents SAHH with no PEG-Mal added (molecular mass of 43 kDa).

To test MEG binding to Cys residues, we pre-incubated SAHH with 1 mM MEG for 1h at 37ºC and monitored SAHH PEGylation. After removing excess of MEG using gel filtration (PD-10 column), the resulting protein was PEGylated with 1:400 or 1:800 protein:PEG-Mal (by mass) for 0 and 10 minutes, at RT, under constant stirring of 140 rpm.

The polyacrylamide gel depicted in Figure 19 shows that in the presence of MEG the relative intensity of the different bands changed, suggesting a change in the number of available Cys residues and indicating that MEG was able to bind to Cys residues, blocking their reactivity towards PEG-Mal.

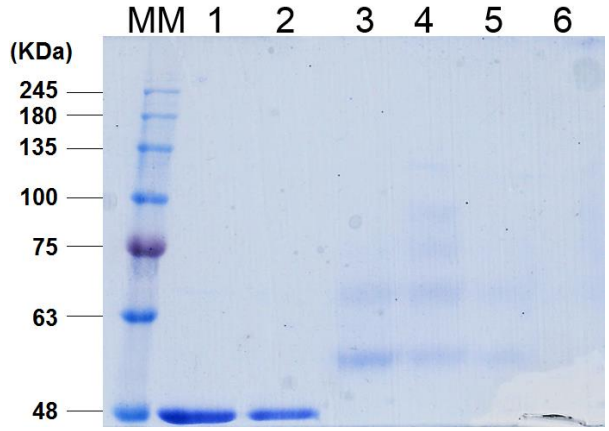

Figure 19 – SDS-PAGE analysis of PEG-Mal binding to S-adenosyl-homocysteine hydrolase (SAHH) after MEG treatment. SAHH was pre-incubated with 1 mM MEG for 1h at 37ºC; after MEG removal, SAHH was incubated with 1:400 or 1:800 protein:PEG-Mal (by mass) at RT for 0 and 10 minutes, under constant stirring of 140 rpm. SAHH with no PEG-Mal added (1, 2); SAHH incubated with 1:400 protein:PEG-Mal (by mass) after 0 and 10 min (3 and 4, respectively); SAHH incubated with 1:800 protein:PEG-Mal (by mass), after 0 and 10 min (5 and 6, respectively).

In order to confirm the results obtained for SAHH PEGylation, assessment of the protein free –SH groups, before and after incubation with MEG followed by protein PEGylation, was performed using the DTNB assay. This compound upon binding to free –SH groups produces a chromophore. The analysis of the results obtained with the DTNB assays indicate a lower level of free –SH groups after MEG incubation and SAHH PEGylation, suggesting that MEG binds to Cys residues in the SAHH protein (data not shown).

In conclusion, the results obtained suggest that the optimized conditions for the PEG-Mal reaction are incubation at RT (25ºC) for 10 min and using a 1:400 SAHH:PEG-Mal ratio (by mass).

4.1.5.2. PEGylation of CBS WT

As stated above, the CBS protein has in its structure 11 Cys residues and as such, 11 bands with MM ranging from 68 to 118 kDa were expected to be observed (68, 73, 78, 83, 88, 93, 98, 103, 108, 113 and 118 kDa). The conditions of the PEGylation reaction optimized for the SAHH protein were then used to monitor CBS PEGylation with and without MEG treatment.

In Figure 20, it is presented the PEGylated CBS WT protein without incubation with MEG (panel A) and with pre-incubation with 1 mM MEG, at 37ºC, during 1h (panel B). In the absence of MEG, five bands above the 63 KDa of the CBS WT protein are visible (Figure 20, panel A), corresponding to the PEG-Mal that bonded to the free Cys residues. It must be noticed that two large bands with MM between 65 KDa–75 KDa and ≈75 KDa were observed, which might correspond to the expected PEGylated forms of 68 and 73 KDa (first band of 65 KDa>MM<75 KDa) and 78, 83 and 88 KDa (second band of ≈75 KDa). A third band near 100 KDa might indicate the presence of the PEGylated forms of 93 and 98 KDa. However, after incubation with MEG, no significant changes were observed to the SDS-PAGE profile. Due to the low resolution of the gel, we were not able to completely exclude a change in the electrophoretic profile in the presence of this aminothiol, and therefore this assay should be repeated using a gel gradient, which presents a higher resolution.

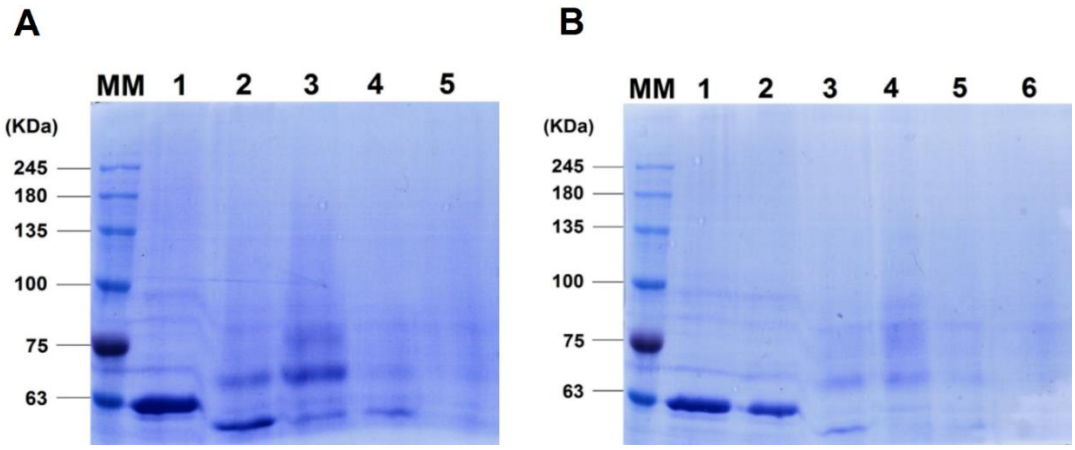

Figure 20 – SDS-PAGE analysis of PEG-Mal binding to cystathionine-β-synthase (CBS) without pre-incubation (Panel A) and after pre-incubation with 1 mM MEG during 1h at 37ºC. Panel A – The CBS WT was incubated with 1:400 or 1:800 protein:PEG-Mal ratio (by mass) at RT for 0 and 10 minutes, under constant stirring of 140 rpm. CBS WT with no PEG-Mal added (1); CBS WT incubated with 1:400 protein:PEG-Mal ratio (by mass) at 0 and 10 min (2 and 3 respectively); CBS WT incubated with 1:800 protein:PEG-Mal ratio (by mass) for 0 and 10 min (4 and 5, respectively). Panel B – CBS WT after incubation with MEG and with no PEG-Mal (1); CBS WT with MEG, after gel filtration (PD MiniTrap) (no PEG-Mal) (2); CBS WT incubated with 1:400 protein:PEG-Mal ratio (by mass) for 0 and 10 min (3 and 4, respectively); CBS WT incubated with 1:800 protein:PEG-Mal ratio (by mass) for 0 and 10 min (5 and 6, respectively).

4.1.6. Enzymatic activity of R>C variants

4.1.6.1. Optimization of an HPLC method for cystathionine quantification

The enzymatic activity of CBS is determined based on the amount of Cth produced in the reaction mixture. Till present, the Cth quantification was based on a GC-FTIR technique after extraction and derivatization using the EZ:faast™ Amino Acid Analysis Kit (Phenomenex). Besides its low sensitivity, this technique is highly expensive. Therefore, an alternative method for detection of Cth using HPLC was optimized based on the technique described by Ravanel S and collaborators (Ravanel et al., 1995).

Using the solvents gradient system as described by the above authors, we performed preliminary HPLC assays to confirm the complete resolution between the amino acids used in the enzymatic reaction mixture, namely the substrates L-Ser and L-Hcy, the reaction

product Cth, and the aminothiol compounds cysteamine and MEG (Figure 21A and B). We also tested the use of norvaline as the internal standard (Figure 21B).

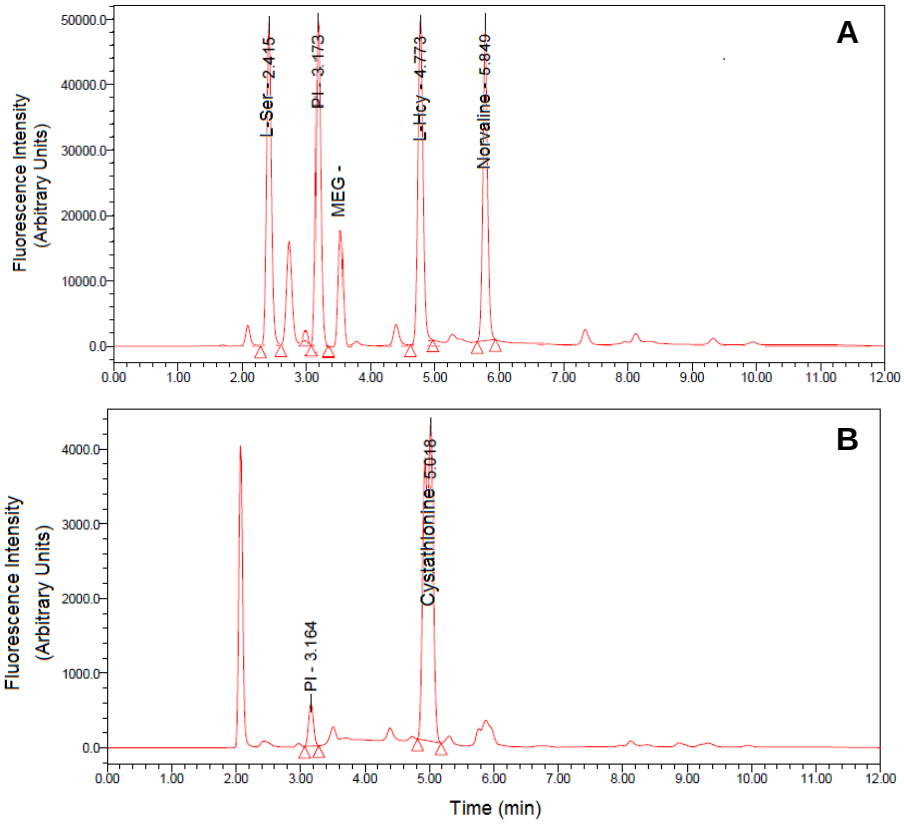

Figure 21 - HPLC chromatograms of standard solutions. (A) L-serine (L-Ser), MEG, norvaline and L-homocysteine (L-Hcy). (B) Cystathionine. (PI) Cysteamine. See text for details.

We found that L-Hcy presented an elution time similar to cystathionine (Figure 21A and B), which could lead to an incorrect quantification of the reaction product. Therefore, we changed the gradient and the running time (15 min), while maintaining the oven temperature, the buffers composition and the flow rate. At time zero, the percentage of buffer B was decreased from 40 to 25%. The initial linear gradient of 25 to 80% solvent B (previously 40-80%) was extended to 8 min (previously 0-6 min), then 80% solvent B was maintained till 11 min and then returned to 25% in 1 min. Using the optimized solvent gradient, a higher resolution of L-Hcy and Cth peaks were obtained (Figure 22).

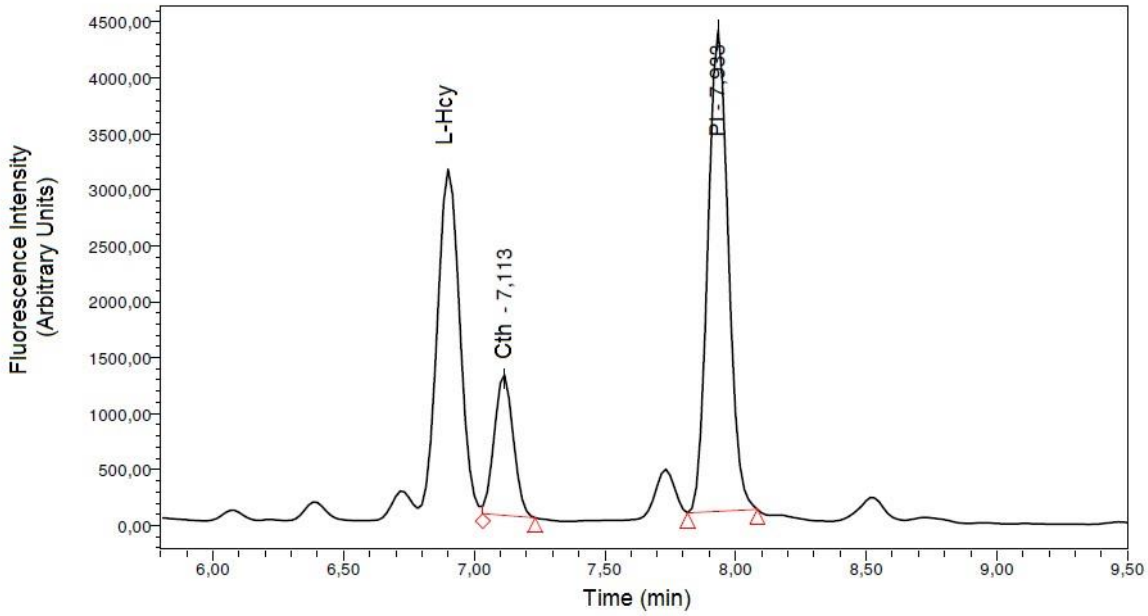

Figure 22 - HPLC chromatogram of the enzymatic reaction of CBS WT. Norvaline was used as the Internal Standard (PI). See text for details.

Using the described conditions, the HPLC quantification of Cth revealed to be linear (R^2 of 0.998) from 7.8 to 500 μM Cth, corresponding to an injected amount of 39 pmol to 2.5 nmol of Cth, as shown in Figure 23.

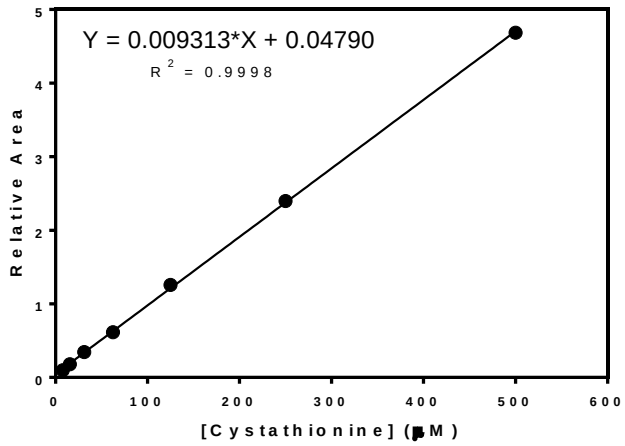

Figure 23 – Calibration curve of L-cystathionine. Standard solutions were measured by HPLC after OPA derivatization and fluorimetric detection. See text for details. The equation of the linear regression line and the corresponding correlation coefficient (R^2) are presented.

4.1.6.2. Enzymatic activity of the R>C CBS variants in the absence and in the presence of MEG

Using the developed HPLC method for Cth quantification, the catalytic activity of the WT and variant proteins were determined. The assays were performed in the absence and after pre-incubation of the proteins with MEG, at 37ºC during 3h. As shown in Figure 24, from the studied variant proteins, only for the p.R121C the HPLC method was unable to detect the produced Cth.

According to our previous studies (Mendes et al., 2015) the p.R336C variant presented around 12% of residual activity. Curiously, for the remaining variants their residual catalytic activity was higher than the WT. These observations were expected for the p.R18C (4.4-fold increase) and for the mutations affecting the R491 residue (R>C and R>H, 9.4- and 6.7-fold increase respectively), but not for the p.R369C variant (2.9-fold increase).

In fact, p.R18C was found with 1% of the allele frequency in normal controls in the Korean population (Lee et al., 2005) and with a residual activity, when expressed in mammalian cells, of 30%. These data leaded to the classification of the p.R18C variant as a single nucleotide polymorphism.

Regarding the mutations affecting the R491 residue, it is well known that amino acid changes in the C-terminal regulatory domain usually lead to CBS proteins presenting a higher residual activity than the WT protein, since they lead to a protein conformation in a pre-activated state (Mendes, Santos et al, 2014).

When compared to the WT CBS, the p.R369C variant presented not only a higher residual activity, but also a similar resistance to proteolysis (Table 5 and Figures 14B and 15). However, its C-terminal domain presented a lower stability (T_{m1}< 5ºC, Table 6).

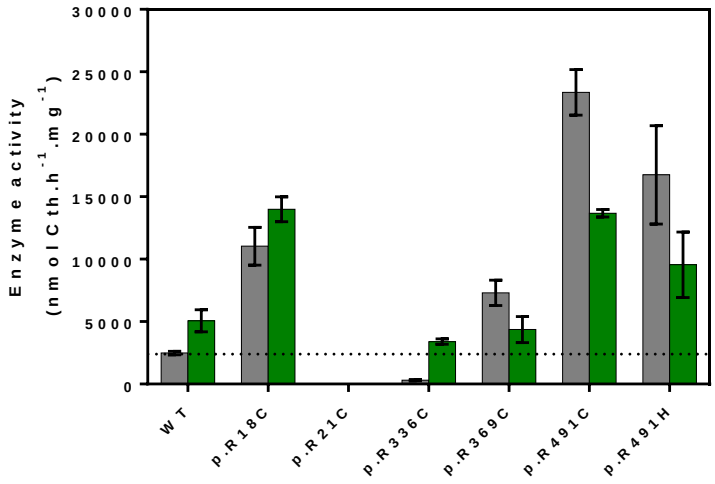

Figure 24 – Specific enzyme activity of CBS recombinant proteins (WT and R>C variants) in the absence (grey columns) and with pre-incubation with 1mM MEG (green columns). The dotted line indicates the catalytic activity for the WT CBS without added thiol compound.

Interestingly, when incubated with 1 mM MEG for 3h, only the p.R18C and p.R336C variants showed an increase in the catalytic activity of 1.3- and 11-fold respectively in respect to the enzymatic activity determined in the absence of the aminothiol compound, which suggests that the potential stabilizing effect of MEG upon binding to R>C may be specific for specific residues.

The variant proteins who presented an increase (or similar) proteolysis rate after pre-incubation with MEG, namely the p.R369C, p.R491C and p.R491H, all showed a 0.6-fold decrease in their catalytic activity, indicating that for these variants the aminothiol does not have an apparent stabilizing effect.

4.2. Characterization of the p.L388P variant identified in a Portuguese CBS deficient patient

The CBS variant L338P was first identified in 2003 in a Spanish patient (Urreizti et al., 2003) with a clinical phenotype of a severe form of classical homocystinuria. In

fibroblasts, the enzyme presented only 3.1% of residual activity and was considered non-responder to Vitamin B_6 (6.4% residual activity in the presence of PLP). Previous characterization of the p.L388P protein in *E. coli* lysates showed a residual enzyme activity lower than 4%. Additionally, native gel experiments allowed to demonstrate this variant did not form tetramers and that monomers were also dramatically diminished, suggesting an intrinsic instability in the tertiary structure (Urreizti et al., 2006).

Recently, our group also identified the p.L338P variant in a Portuguese classical homocystinuric patient. This patient has been followed in the Metabolic Unit of Hospital Santa Maria, and the clinical phenotype after Vitamin B6 supplementation was suggestive of a responsive patient. In order to elucidate the observed clinical evaluation, we initiated the expression, purification and characterization of the p.L338P variant.

4.2.1. Expression and purification of recombinant WT and p.L388P CBS variant

Using site directed mutagenesis, the c.1013T>C mutation (p.L338P) was successfully introduced in the expression construct pET28b-6xHis-pepT-hCBSWT. DNA sequencing allowed confirming the presence of the desired mutation (Figure 25) and the absence of additional DNA changes.

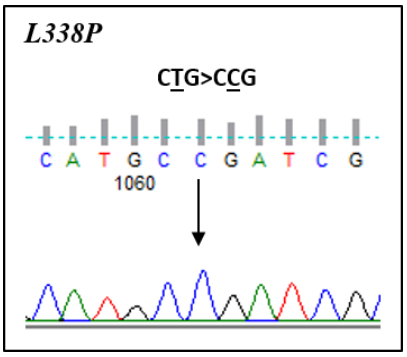

Figure 25 – Partial sequence panel of the CBS cDNA after site-directed mutagenesis for the introduction in the CBS cDNA of the c.1013T>C mutation encoding for the p.L388P variant of CBS protein. The arrow indicates the position of the single base substitution in the cDNA sequence. The wild-type and mutated codon are shown with the base change underlined.

Using the previously described expression conditions, the recombinant p.L388P variant protein was eluted from the Ni-NTA column with 500 mM imidazole (Figure 26).

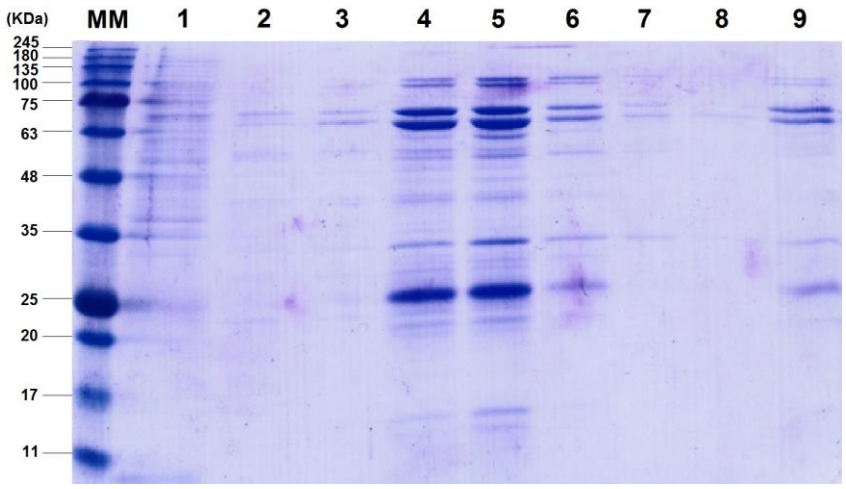

Figure 26 – SDS-PAGE analysis of the affinity chromatographic purification of recombinant p.L338P variant. Recombinant protein was purified by immobilized metal affinity chromatography using a gradient of imidazole concentration: 10 mL of 20 mM (1), 10 mL of 50 mM (2), 2.5 mL of 75 mM (3), and 5 x 500 µL of 500 mM Imidazole (4-8). (MM) Pre-stained protein ladder. (9) Purified CBS protein after removal of imidazole by gel filtration (PD-10 column). In each lane an aliquot of 16 µL of each imidazole elution fraction was applied.

As shown in Figure 26, the p.L338P variant was produced in low levels (59.17 µg/L culture vs. 597.94 µg/L culture for WT) and with a purity grade of only ≈4% (vs. 80% for WT), which precluded its use for proteolysis and DSF assays. The low expression levels indicate that the protein will probably be found mainly in the insoluble fraction of the *E.coli* lysate, thus pointing to a protein with a high propensity to aggregate.

4.6.4. Effect of PLP on the catalytic activity of the p.L338P variant

The low yield obtained for the p.L338P variant already anticipated a variant protein with reduced enzymatic activity. In fact, with the developed HPLC method we were not able to detect the Cth produced by this variant enzyme, nor in the absence nor in the presence of

PLP (Table 7). Interestingly, the WT protein did not significantly change its activity in the presence of PLP, as previously reported by our group (Mendes, Santos et al., 2014).

Table 7 – Residual enzyme activity of the WT and the p.L338P CBS proteins in the absence and presence of PLP

	Specific enzyme activity (nmol $Cth.h^{-1}.mg^{-1}$)	
	-PLP	PLP (2.7 mM)
WT	100	92
p.L338P	n.d.	n.d.

Note: the enzymes were pre-incubated at 37ºC, for 1h, in the absence and in the presence of 1 mM PLP. (n.d.) not detected.

The low yield of the p.L338P variant did not allow the biochemical and structural characterization of this protein. Changes in the expression procedure should be implemented (higher culture volume and/or increased expression time) in order to obtain the variant protein with a good purity grade and in high amounts to perform the characterization studies.

5. General Discussion

In this work, we studied the effect of two aminothiol compounds, CySH and MEG, on the partial repair of five CBS variants carrying an R>C substitution: p.R18C, p.R121C, p.R336C, p.R369C and p.R491C. Until today, similar studies were implemented for CySH using other proteins than CBS, but with MEG only the p.R336C has been investigated (Mendes et al., 2015). To study whether the observed effects of incubation with CySH and MEG would not result in unwanted interactions with other Cys residues in CBS, we included in our work the CBS WT. We also extended our studies to the p.R491H variant as a negative control, to determine if the effect of the aminothiol compounds was specific for the variant Cys residue (Cys-specific). The rescue by aminothiols was assessed by monitoring not only CBS's function, but also some of its biophysical properties, such as thermal stability and conformational flexibility. Finally, a method was developed to tentatively confirm MEG binding to the CBS variants.

From the introduced changes, the p.R18C has been considered as a polymorphism (Lee et al, 2005), the p.R121C, p.R336 and p.R491C variants have been categorized as disease-causing, and the p.R164C has never been identified among the classical homocystinuric population. According to our data, the p.R18C, p.R369C, p.R491C and p.R491H variants presented enzymatic activities in the range or even higher than the WT protein and, except for p.R18C, all of the variants showed a proteolysis rate similar to the WT protein. Although the p.R18C had never been expressed in *E. coli* and characterized *in vitro*, it is important to highlight that this behavior has also been found for other CBS variants affecting the N-terminal CBS domain. In fact, when compared to the WT form, the p.P49L and p.P78R variants identified in CBS deficient patients present similar (p.P49L) or even higher (p.P78R) specific enzyme activities and proteolysis rate than the WT (Hnízda et al., 2012). These observations corroborate previous data obtained by our group were a mild effect of amino acid changes occurring within the N-terminal domain were postulated (Mendes et al., 2014).

Likewise, the variant proteins presenting amino acid changes in the C-terminal regulatory domain (p.R491C and p.R491H) presented specific enzyme activities higher than the WT. This is a known phenomenon resulting from a higher accessibility of the substrates to the catalytic domain upon conformational changes in the C-terminal domain inducing a

pre-activated structure. As shown in Figure 27, no major interactions will be disrupted by the substitution of Arg by a Cys.

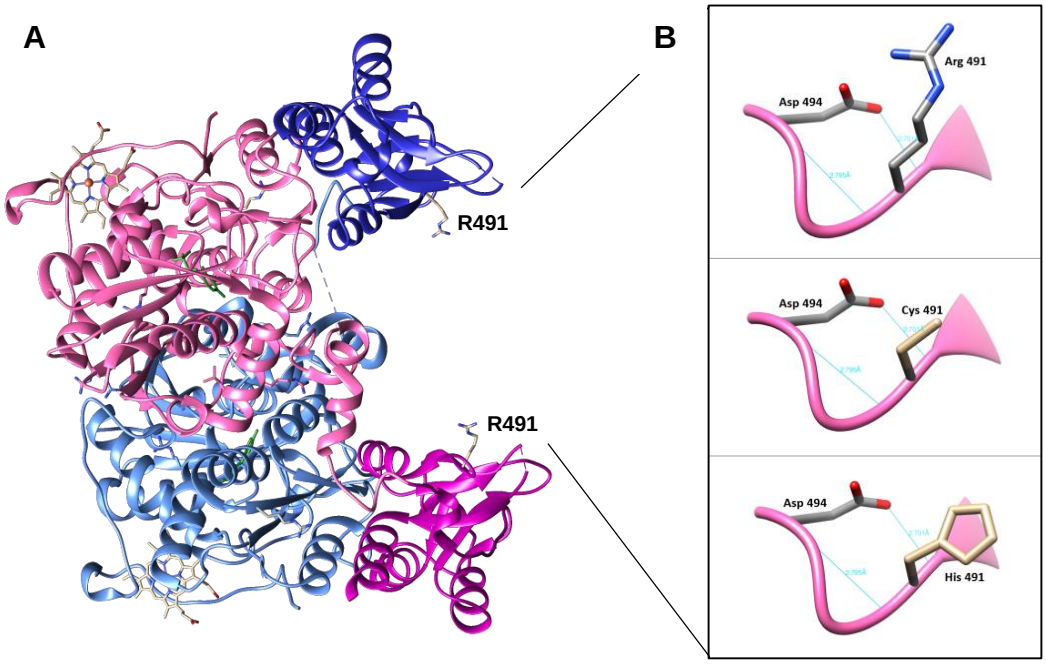

Figure 27 – Structural representation of a CBS dimer with the R491 residue highlighted. In **A**, chain A is represented in pink and chain B is represented in blue with the corresponding C-terminal catalytic domains in dark pink and dark blue, respectively. In **B** the established interactions of the WT residue and the R>C and R>H substitutions are shown. Images were obtained using the crystal structure of a truncated form of CBS (PDB ID 4COO) and the UCSF Chimera, version 1.8.1.

Concerning the p.R369C variant, contradictory data can be found in the literature. In fact, when studied in crude extracts of *E. coli* lysates, this variant presents only 15% of the WT activity and low expression (Kozich et al., 2010). However, when studied in CHO cells, the p.R369C presented ≈70% of residual activity, leading the authors to state that the net effect of the p.R369C substitution on quaternary structure and activity appears to be moderate in mammalian cells (Janošík et al., 2009), and that from the clinical standpoint may represent a functionally neutral genetic variant. Taking into consideration the structure presented in Figure 28 and the ASA value of 21% (partial buried residue) calculated for this residue, no significant structural changes would be predicted.

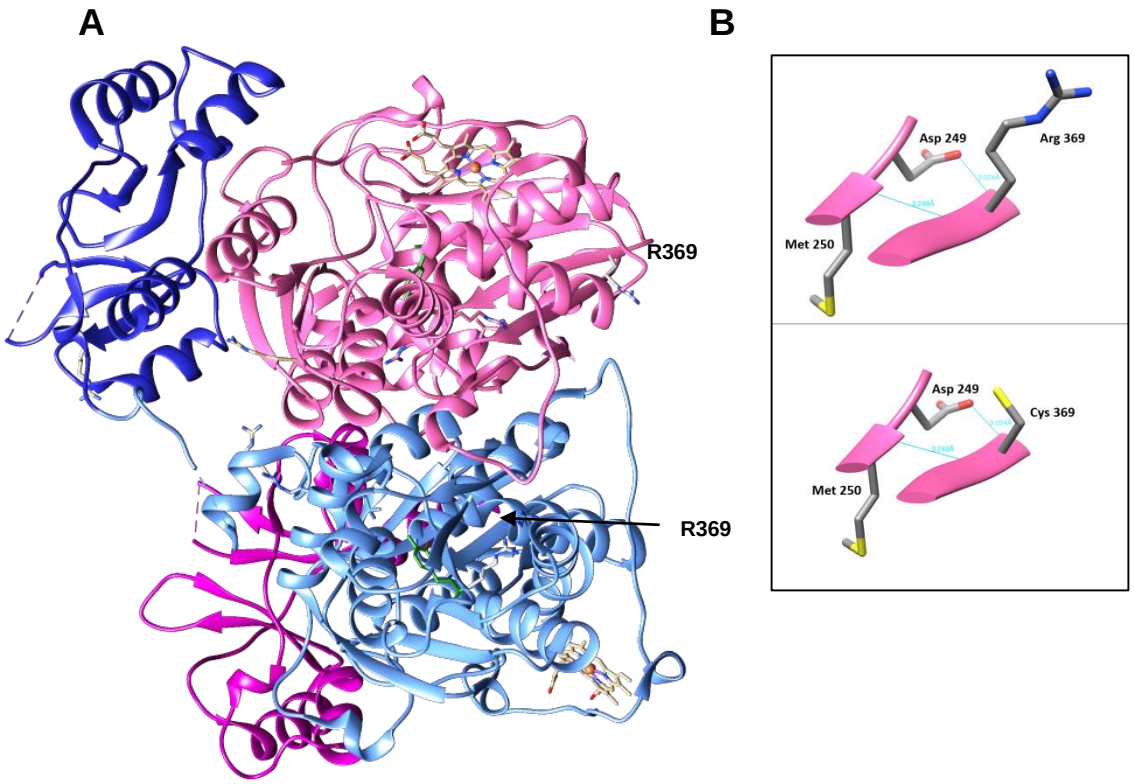

Figure 28 - Structural representation of a CBS dimer with the R369 residue highlighted. In **A**, chain A is represented in pink and chain B is represented in blue with the corresponding C-terminal catalytic domains in dark pink and dark blue, respectively. In **B** the established interactions of the WT residue and the R>C substitutions are shown. Images were obtained using the crystal structure of a truncated form of CBS (PDB ID 4COO) and the UCSF Chimera, version 1.8.1.

As shown in Figure 29, the substitution of an Arg by a Cys in residues 121 and 164 will probably lead to dramatic changes in the interactions that those residues establish with Glu239, Thr87 and Glu110 (for Arg121) or Glu187 (for Arg164), which would explain the low expression levels obtained for both p.R121C and p.R164C variants.

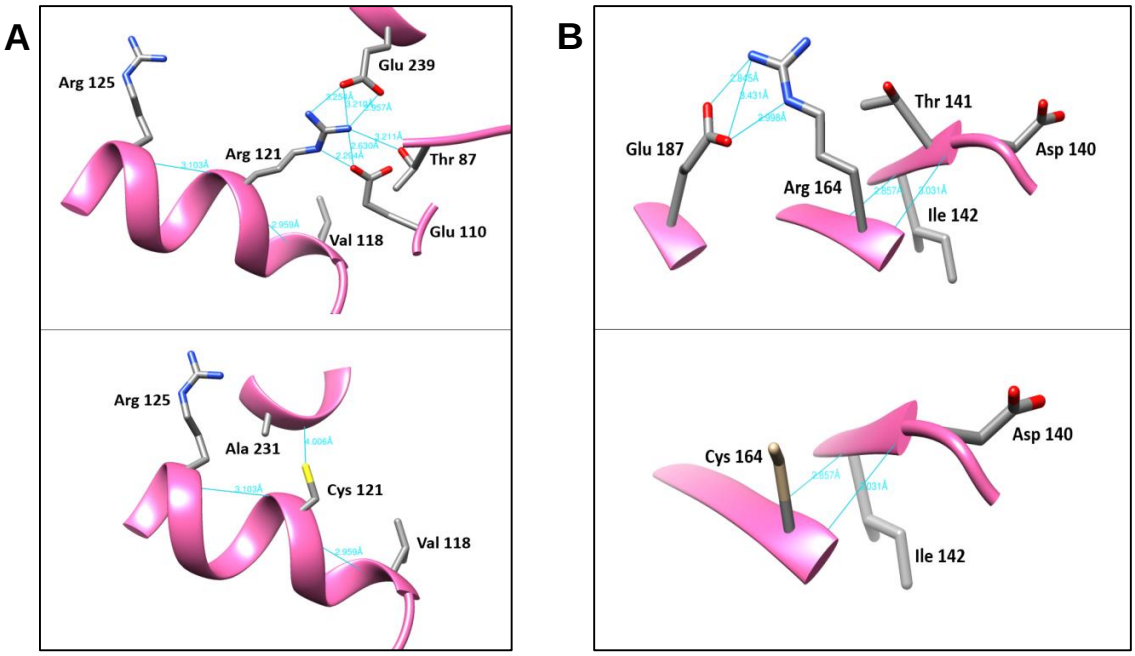

Figure 29 - Structural representation of the R121 (**A**) and R164 (**B**) residues. The interactions established by the WT residue and by the R>C substitutions are shown. Images were obtained using the crystal structure of a truncated form of CBS (PDB ID 4COO) and the UCSF Chimera, version 1.8.1.

Kožich and colleagues (Kožich et al., 2010) analyzed the role of solvent accessibility of various mutants as a determinant of their folding and activity. In general, buried mutations exhibited much more profound detrimental effects on enzyme folding and activity compared to solvent-accessible mutations (Kožich et al., 2010), showing that the topology of mutations predicts in part the behavior of variant CBS proteins, and that misfolding may be an important and frequent pathogenic mechanism in CBS deficiency. This states the importance of solvent accessibility of mutations rather than their location in enzyme domains in predicting the properties of mutant enzymes. The accessible surface area (ASA) describes the area over which contacts between protein and solvent occur. The extent to which an amino acid residue interacts with its environment, the solvent and the protein core, is naturally proportional to the degree to which it is exposed to these environment. As such, to study the rescuing effect of aminothiols compounds, residues with different ASA values were selected.

Previous work of our group already showed that when compared to MEG, CySH has a much more attenuated effect over the p.R336C variant (Mendes et al., 2015). We now show that this effect is also observed for the other R>C variants included in this study. In fact,

CySH increased the proteolysis rate, not only of the majority of the R>C variants, but also of the WT protein. These data suggest that in the presence of this aminothiol compound, the proteins present a higher conformational flexibility and/or a more opened structure. Only the p.R18C (proposed high ASA value) and p.R121C (buried residue, ASA value of 6%) were less prone to proteolysis by trypsin. As such, CySH does not present a universal Cys-rescuing effect.

Concerning the aminothiol compound MEG, a non-universal Cys rescue effect was also found. In fact, in the presence of this compound only the p.R18C, p.R121C and p.R336C variants decreased their proteolysis rate, which is translated into a higher catalytic activity. Again, the ASA values of these residues are quite different, since they range from highly exposed (p.R18C), to partially exposed (p.R336C; ASA 21%) and buried (p.R121C; ASA 6%) residues. The DSF studies performed indicate that for the p.R336C variant a stabilization of the catalytic domain may explain the observed effect.

Taken together, our data strongly suggest that the rescue of R>C variants by aminothiol compounds is not an universal effect, but it will rather depend on the affected residue, and as such, a case-by-case study should always be performed.

In order to monitor the number of available Cys residues accessible to thiol modification, a PEG-Mal assay was established. Although the obtained data did not allow to clearly show the expected changes in the MM, we postulate that further optimization of this technique will contribute to identify the presence of mixed thiols.

During this work, a CBS deficient patient was identified presenting the c.1013T>C mutation (p.L338P). As stated before, around 50% of CBS deficient patients respond to high doses of pyridoxine, thus being classified Vitamin B_6-responsivess. Therefore, we embraced the expression and characterization of the p.L388P variant and, in particular, the enzyme response to the presence of PLP. The low expression levels of the produced variant precluded the complete biochemical and biophysical characterization, but suggest a strong effect on the protein folding, as would be expected by a change of a Leu to a Pro residue. The residue L338 lies at the core of the catalytic domain, at the end of helix12, and a change to a Pro probably will lead to a break in this secondary structure element. Work is on progress to increase the amount of expressed and purified variant in order to allow its thorough characterization.

6. References

Aitken S, Kirsch J. The enzymology of cystathionine biosynthesis; strategies for the control of substrate and reaction specificity. *Archives of Biochemistry and Biophysics* 2005; 433: 166-175.

Aly A, Arai M, Hoyer L. Cysteamine Enhances the Procoagulant Activity of Factor Vil-East Hartford, A Dysfunctional Protein due to a Light Chain Thrombin Cleavage Site Mutation (Arginine-1689 to Cysteine). *Journal of Clinical Investigation* 1992; 89: 1375-1381.

Banerjee R, Zou C. Redox regulation and reaction mechanism of human cystathionine-b-synthase: a PLP-dependent hemesensor protein. *Archives of Biochemistry and Biophysics* 2005; 433: 144-156.

Bashir H, Dekair L, Mahmoud Y, Ben-Omran T. Neurodevelopmental and Cognitive Outcomes of Classical Homocystinuria: Experience from Qatar. *JIMD Reports* 2015; 21: 89–95.

Borges C, Sherma N. Techniques for the Analysis of Cysteine Sulfhydryls and Oxidative Protein Folding. *Antioxidants & Redox Signaling* 2014; 21: 511-531.

Bradford MM. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Analytical Biochemistry* 1976; 72: 248-254.

Cantoni G. The Nature of the Active Methyl Donor Formed Enzymatically from L-Methionine and Adenosinetriphosphate. *J Am Chem Soc* 1952; 74: 2942–2943.

Carmel R., Jacobsen DW. Homocysteine. In: *Health and Disease*. Cambridge, Cambridge University Press, 2001.

Coulondre C, Miller J, Farabaugh P, Gilbert W. Molecular basis of base substitution hotspots in Escherichia coli. *Nature* 1978; 274: 775-780.

Cruysberg J, Boers G, Trijbels J, Deutman A. Delay in dignosis of homocystinuria: retrospective study of consecutive patients. *Br Med J* 1996; 313: 1037-1040.

Cubtrecasas P, Anfinsen C. Methods in enzymology. New York, Academic Press, 1968.

De Franchis R, Kraus E, Kozich V, Sebastio G, Kraus J. Four novel mutations in the cystathionine beta-synthase gene: effect of a second linked mutation on the severity of the homocystinuric phenotype. *Hum Mutat* 1999; 13: 453-457.

Ellison D, Beynon RJ, Hinton J, Hubbard SJ. Limited proteolysis of native proteins: The interaction between avidin and proteinase K. *Protein Science* 1995; 4: 1337-1345.

Di Leandro L, Maras B, Schinina M, Dupre S, Koutris I, Martin F, et al. Cystamine restores GSTA3 levels in Vanin-1 null mice. *Free Radic Biol Med* 2008; 44: 1088–1096.

El-Said M, Badii R, Bessisso M, Shahbek N, El-Ali M, et al. A Common Mutation in the CBS Gene Explains a High Incidence of Homocystinuria in the Qatari Population. *Human Mutation* 2006; 27: 719-722.

Ellman GL. Tissue sulfhydryl groups. *Archives of Biochemistry and Biophysics* 1959; 82:70–77.

Finkelstein J. Methionine metabolism in mammals. *J Nutr Biochem* 1990; 1: 228-237.

Finkelstein J, Martin J. Methionine metabolism in mammals. Distribution of homocysteine between competing pathways. *J Biol Chem* 1984; 259: 9508-9513.

Finkelstein J, Martin J. Homocysteine. *The International Journal of Biochemistry & Cell Biology* 2000; 32: 385-389.

Fontana A, Laureto P, Spolaore B, Frare E. Probing protein structure by limited proteolysis. *Acta Biochimica Polonica* 2004; 51: 299–321.

Fried MW et al. PEGinterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. *New England Journal of Medicine* 2002; 347: 975–982.

Gaustadnes M, Wilcken B, Oliveriusova J, McGill J, Fletcher J, Kraus J, Wilcken D. The molecular basis of cystathionine beta-synthase deficiency in Australian patients: genotype-phenotype correlations and response to treatment. *Hum Mutat* 2002; 20: 117-126.

Guidolin D, Agnati LF, Albertin G, Tortorella C, Fuxe K. Bioinformatics aggregation predictors in the study of protein conformational diseases of the human nervous system. *Electrophoresis* 2012; 33: 3669-3679.

Hadziyannis SJ et al. Peginterferon-alpha2a and ribavirin combination therapy in chronic hepatitis C: a randomized study of treatment duration and ribavirin dose. *Annals of Internal Medicine* 2004; 140: 346–355.

Hnízda A, Jurga V, Raková K, Kožich V. Cystathionine beta-synthase mutants exhibit changes in protein unfolding: conformational analysis of misfolded variants in crude cell extracts. *J Inherit Metab Dis* 2012; 35: 469–477.

Horie Y, Fazio S, Westerlund J, Weisgraber K,. Rall S. The Functional Characteristics of a Human Apolipoprotein E Variant (Cysteine at Residue1 42) May Explain Its Association with Dominant Expression of Type III Hyperlipoproteinemia. *The Journal of Biological Chemistry* 1992; 267: 1962-1968.

Hu F, Gu Z, Kozich V, Kraus J, Ramesh V, Shih V. Molecular basis of cystathionine beta-synthase deficiency in pyridoxine responsive and nonresponsive homocystinuria. *Hum Mol Genet* 1993; 2: 1857- 1860.

Hubbard SJ. The structural aspects of limited proteolysis of native proteins. *Biochim Biophys Acta* 1998; 1382: 191-206.

Janošík M, Sokolová J, Janošíková B, Krijt J, Klatovská V, Kožich V. Birth Prevalence of Homocystinuria in Central Europe: Frequency and Pathogenicity of Mutation c.1105C>T (p.R369C) in the Cystathionine Beta-Synthase Gene. *J Pediatr* 2009; 154: 431-437.

Kang Z, Zhang J, Zhou J, Qi Q, Du G, Chen J. Recent advances in microbial production of δ-aminolevulinic acid and vitamin B12. *Biotechnology Advances* 2012; 30:1533-1542.

Kery V, Bukovska G, Kraus J. Transsulfuration depends on heme in addition to pyridoxal 5'-phosphate. Cystathionine ß-synthase is a heme protein. *J Biol Chem* 1994; 269: 25283-25288.

Khan S, Vihinen M. Spectrum of disease-causing mutations in protein secondary structures. *BMC Structural Biology* 2007; 7: 1-18.

Kopecká J, Krijt J, Raková K. Restoring assembly and activity of cystathionine β-synthase mutants by ligands and chemical chaperones. *J Inherit Metab Dis* 2011; 34: 39-48.

Kožich V, Kraus J. Screening for mutations by expressing patient cDNA segments in *E. coli* - homocystinuria due to cystathionine ß-synthase deficiency. *Hum. Mutation* 1992; 1: 113-123.

Kožich V, Kraus J. Homocysteine in health and disease. Cambridge: Cambridge University Press, 2011.

Kožich V, Sokolová J, Klatovská V, Krijt J, Janošik M, Jelínek K, Kraus J. Cystathionine b-Synthase Mutations: Effect of Mutation Topology on Folding and Activity. *Human Mutation* 2010; 31: 809–819.

Kraus J. Biochemistry and molecular genetics of cystathionine β-synthase deficiency. *European Journal of Pediatrics* 1998; 157: S50-S53.

Kraus J, Janosik M, Kozich V, Mandell R, Shih V, Sperandeo M, et al. Cystathionine beta-synthase mutations in homocystinuria. *Hum Mutat* 1999; 13: 362-375.

Kraus J, Oliveriusová J, Sokolová J, Kraus E, Vlcek C, de Franchis R, Maclean KN, Bao L, Bukovsk, Patterson D, Paces V, Ansorge W, Kozich V. The human cystathionine beta-synthase (CBS) gene: complete sequence, alternative splicing, and polymorphisms. *Genomics* (1998); 52: 312-324.

Kunkel T. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proceedings of the National Academy of Sciences USA* 1985; 82: 488-492.

Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 1970; 227: 680-685.

Leandro P, Gomes C. Protein misfolding in conformational disorders: rescue of folding defects and chemical chaperoning. *Mini-Reviews in Medicinal Chemistry* 2008; 8: 901-911.

Lee S, Lee D, Yoo H, Koo S, Park E, Park J, et al. Identification and functional analysis of cystathionine beta-synthase gene mutations in patients with homocystinuria. *J Hum Genet* 2005; 50: 648-654.

Lesort M, Lee M, Tucholski J, Johnson G. Cystamine inhibits caspase activity. Implications for the treatment of polyglutamine disorders. *J Biol Chem* 2003; 278: 3825–3830.

Lohinai Z, Benedek P, Fehér E, GyoÈr A, Rosivall L, Fazekas Á, et al. Protective effects of mercaptoethylguanidine, a selective inhibitor of inducible nitric oxide synthase, in ligature-induced periodontitis in the rat. *British Journal of Pharmacology* 1998; 123: 353-360.

Lu J, Deutsch C. Pegylation: a method for assessing the topology accessibility in Kv1.3. *Biochemistry* 2001; 40: 13288-13301.

Majtan T, Singh LR, Wang L, Kruger WD, Kraus JP. Active cystathionine beta-synthase can be expressed in heme-free systems in the presence of metal-substituted porphyrins or a chemical chaperone. *Journal of Biological Chemistry* 2008; 283: 34588-95.

Malvern. Dynamic Light Scattering (DLS). Retrieved from Malvern: http://www.malvern.com/en/products/technology/dynamic-light-scattering/ (consulted in: 01-07-2014).

McKeever W, Weir D, Molloy A, Scott J. Betaine-homocysteino methyltransferase: Organ distribution in man, pig and rat and subcellular distribution in the rat. *Clin Sci* 1991; 81: 551-556.

Meier M, Janosik M, Kery V, Kraus J, Burkhard P. Structure of human cystathionine beta-synthase: a unique pyridoxal 5'-phosphate-dependent heme protein. *Embo J* 2001; 20: 3910-3916.

Meier M, Janosik M, Kery V, Kraus J, Burkhard P. Structure of human cystathionine beta-synthase: a unique pyridoxal 5'-phosphate-dependent heme protein. *Embo J* 2001; 20: 3910-3916.

Meier M, Oliveriusova J, Kraus J Structural insights into mutations of cystathionine beta-synthase. *Biochim Biophys Acta* 2003; 1647: 206-213.

Melenovská P, Kopecká J, Krijt J, Hnízda A, Raková K, Janosík M, Wilcken B, Kozich V. Chaperone therapy for homocystinuria: the rescue of CBS mutations by heme arginate. *Journal of Inherited Metabolic Diseases* 2015; 38: 287-294.

Mendes M, Colaço H, Smith D, Ramos R, Pop A, Dooren S, Almeida I, Kluijtmans L, Janssen M, Rivera I, Salomons G, Leandro P, Blom H. Reduced response of Cystathionine Beta-Synthase (CBS) to S-Adenosylmethionine (SAM): Identification and functional analysis of CBS gene mutations in Homocystinuria patients. *Journal of Inherited Metabolic Diseases* 2014; 37:245-254.

Mendes M, Santos A, Smith D, Lino P, Colaço H, Almeida I, Vicente J, Salomons G, Rivera I, Blom H, Leandro P. Insights into the Regulatory Domain of Cystathionine Beta-Synthase: Characterization of Six Variant Proteins. *Human Mutation* 2014; 35: 1195-1202.

Mendes M, Smith D, Colaco H, Santos S, Rivera I, Ben-Omran T, et al. A new approach on protein folding correction: rescue of arginine to cysteine mutations using thiol compounds. *FEBS Journal* 2013; 280: 114.

Mendes M, Smith D, Vicente J, Almeida I, Ben-Omran T, Salomons G, Rivera I, Leandro P, Blom H. Small aminothiol compounds improve the function of Arg to Cys variant proteins: effect on the human cystathionine β-synthase p.R336C. *Human Molecular Genetics* 2015 (in press).

Mendre C, Mouillac B. Chaperons pharmacologiques: Un espoir thérapeutique pour les pathologies conformationnelles. *Medecine Sciences* 2010; 26: 627-635.

Mudd S, Levy H, Kraus J. Disorders of transsulfuration. In: Scriver CR, Beaudet AL, Sly WS, Valle D, editors. The Metabolic and Molecular Bases of Inherited Disease. New York: McGraw-Hill, 2001: 2007-2056.

Mudd S, Levy H, Skovby F. Disorders of transsulfuration. In: The Metabolic Basis of Inherited Disease. Scriver CR, Beaudet AL, Sly WS, Valle D, 6 ed. New York: McGraw-Hill, Inc, 1989: 1279-1327.

Mudd S, Skovby F, Levy H, Pettigrew K, Wilcken B, Pyeritz R, et al. The natural history of homocystinuria due to cystathionine beta-synthase deficiency. *Am J Hum Genet* 1985; 37: 1-31.

Münke M, Kraus J, Ohura T, Francke U. The gene for cystathionine ß-synthase (CBS) maps to the subtelomeric region on human chromosome 21q and to proximal mouse chromosome 17. *Am J Hum Genet* 1988; 42: 550-559.

Niesen F, Berglund H, Vedadi M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature Protocols* 2007; 2: 2212-2221.

Perlmutter D. Chemical chaperones: a pharmacological strategy for disorders of protein folding and trafficking. *Pediatr Res* 2002; 52: 832–836.

Peterson T, Brown I. Cysteamine in combination with N-acetylcysteine prevents acetaminophen-induced hepatotoxicity. *Can J Physiol Pharmacol* 1992; 70: 20-28.

Pey A, Maltan T, Sanchez-Ruiz J, Kraus J. Human cystathionine β-synthase (CBS) contains two classes of binding sites for S-adenosylmethionine (SAM): complex regulation of CBS activity and stability by SAM. *Biochem J* 2013; 449: 109-121.

Pillai A, Veeranan-Karmegam R, Dhandapani K, Mahadik S. Cystamine prevents haloperidol-induced decrease of BDNF/TrkB signaling in mouse frontal cortex. *J Neurochem* 2008; 107: 941–951.

Plasmid map of pET-28_b_(+). Retrieved from: http://www.biovisualtech.com/bvplasmid/pET-28_b_(+).htm (consulted in: 01-07-2014).

Prescott L, Sutherland G, Park J, Smith I, Proudfoot A. Cysteamine, methionine, and penicillamine in the treatment of paracetamol poisoning. *Lancet* 1976; 2: 109-113.

Ravanel S, Droux M, Douce R. Methionine biosynthesis in higher plants. I. Purification and characterization of cystathionine gamma-synthase from spinach chloroplasts. *Archives of Biochemistry and Biophysics* 1995; 316: 572-584.

Refsum H, Fredriksen A, Meyer K, Ueland P, et al. Birth prevalence of homocystinuria. *J Pediatr* 2004; 144: 830-832.

Reid E. Organic Chemistry of Bivalent Sulfur. 1, New York: Chemical Publishing Company, Inc, 1958.

Robyt JF, Ackerman RJ, Chittenden CG. Reaction of protein disulfide groups with Ellman's reagent: a case study of the number of sulfhydryl and disulfide 27 groups in Aspergillus oryzae -amylase, papain, and lysozyme. *Arch Biochem Biophys* 1971; 147: 262-269.

Sartor M. Dynamis Light Scattering to determine the radius of small beads in Brownian motion in a solution. San Diego: University of California, 2010.

Schwartz E, Shapiro B. The Protection of Mice against Radiation by 2-Mercaptoethylguanidine and Its Disulfide. *Radiation Research* 1960; 13: 768-775.

Sedlak J, Lindsay RH. Estimation of total, protein-bound, and nonprotein sulfhydryl groups in tissue with Ellman's reagent. *Anal Biochem* 1968; 25: 192-205.

Singh L, Chen X, Kozich V, Kruger W. Chemical chaperone rescue of mutant human cystathionine b-synthase. *Mol Genet Metab* 2007; 91: 335-342.

Skovby F, Kraus J. The homocystinurias. In: Connective Tissue and its Heritable Disorders: Molecular, Genetic, and Medical Aspects, Royce PM, Steinmann B, vol in press. New York: Wiley-Liss, 1999.

Smith D, Mendes M, Kluijtmans L, Janssen M, Smulders Y, Blom H. A liquid chromatography mass spectrometry method for the measurement of cystathionine β-synthase activity in cell extracts. *J Chromatogr B Analyt Technol Biomed Life Sci* 2012; 91: 186-191.

Smyth D et al. Reactions of N-ethylmaleimide. *Journal of the American Chemical Society* 1960; 82: 4600-4604.

Stauffer M. Chazin Lab. Retrieved from Chazin Lab Protocol for Limited Proteolysis: http://structbio.vanderbilt.edu/chazin/wisdom/labpro/proteolysis.html (consulted in: 26-06-2014).

Steward R, MacArthur M, Laskowski R, Thorto J. Molecular basis of inherited diseases: a structural perspective. *Trends in Genetics* 2003; 19: 505-513.

Talal AH et al. Pharmacodynamics of PEG-IFN alpha differentiate HIV/HCV coinfected sustained virological responders from nonresponders. *Hepatology* 2006; 43: 943–953.

Taoka S, Widjaja L, et al. Assignment of enzymatic functions to specific regions of the PLP-dependent heme protein cystathionine beta-synthase. *Biochemistry* 1999; 38: 13155-13161.

Urreizti R, Asteggiano C, Cozar M, Frank N, Vilaseca M, Grinberg D, Balcells S. Functional Assays Testing Pathogenicity of 14 Cystathionine-Beta Synthase Mutations. *Human Mutation* 2006; 27: 211-219.

Urreizti R, Balcells S, Rodés M, Vilarinho L, Baldellou A, Couce M, Muñoz C, Campistol J, Pintó X, Vilaseca M, Grinberg D. Spectrum of CBS mutations in 16 homocystinuric patients from the Iberian Peninsula: high prevalence of T191M and absence of I278T or G307S. *Human Mutation* 2003; 22: 103-107.

Vitkup D, Sander C, Church G (2003). The amino-acid mutational spectrum of human genetic disease. *Genome Biology* 2003; 4: R72.

Yamanishi M, Kabil O, Sen S, Banerjee R. Structural insights into pathogenic mutations in heme-dependent cystathionine-beta-synthase. *Journal of Inorganic Biochemistry* 2006; 100: 1988–1995.

Zhang W, He H, Wang H, Wang S, Li X, Liu Y, Jiang H, Jiang H, Yan Y, Wang Y, Liu X. Activation of transsulfuration pathway by salvianolic acid a treatment: a homocysteine-lowering approach with beneficial effects on redox homeostasis in high-fat diet-induced hyperlipidemic rats. *Nutrition & Metabolism* 2013; 10: 68-79.

Agradecimentos

Todo o trabalho realizado para a concretização desta tese de mestrado não teria sido possível sem a preciosa ajuda e apoio de um grupo de pessoas muito importantes. Quer a nível profissional, quer a nível pessoal, estas pessoas foram e são valiosas para mim, e a elas quero desde já agradecer profundamente.

Gostaria de agradecer primeiramente à minha orientadora, Professora Paula Leandro, por me dar a oportunidade de trabalhar no grupo de Metabolismo e Genética (MetGen para os amigos) e pela sua fundamental orientação, por toda a paciência, motivação, disponibilidade, ensinamentos e atenção. Obrigada por ter acreditado no meu potencial, e espero ter correspondido às suas expectativas.

À minha co-orientadora, Professora Fátima Ventura, que sempre se mostrou disponível para todas as minhas dúvidas e com o seu "olho clínico" tanto ajudou na escrita desta tese, gostaria de agradecer pela sua dedicação, partilha de conhecimento e por todo o constante acompanhamento, ajuda e exigência na co-orientação do meu percurso nesta tese de mestrado.

Aos investigadores que me apoiaram no laboratório quando já me era permitido "fazer experiências a solo", sempre disponíveis para qualquer dúvida que tivesse, João Vicente e João Leandro, obrigada pela paciência, ensinamentos e disponibilidade.

Às senhoras do laboratório, Ana e dona Lurdes, obrigada pelos conselhos que sempre tinham para me dar, incentivo, pelas histórias e gargalhadas partilhadas.

Ao meu grupo de "amiguinhos do laboratório", Hana, Sofia, Mariana, Cristina, Ana, Israel, Marco e Paulo, obrigada pelos belos almoços, de partilha de histórias e gargalhadas. Um obrigada especial à Hana, que sempre se mostrou disponível para todas as dúvidas que me surgiam. Um obrigada especial também ao Israel, pela ajuda preciosa nas últimas experiências com o HPLC. O trabalho com pessoas de quem gostamos é sempre mais prazeroso.

Às minhas "velhas" amigas da minha querida cidade de Leiria, amigas desde o 9º ano, Mónica e Diana, obrigada por permanecerem sempre verdadeiras, e verdadeiras amigas há tantos anos. Pelos desabafos que tantas vezes tiveram de ouvir das "desgraças engraçadas"

que por vezes acontecem e pelo apoio nas situações menos boas, sempre vamos dando umas boas gargalhadas.

À minha amiga desde o 5º ano, Irina, obrigada por seres a amiga de sempre, ainda que não falemos tão regularmente, quando estamos juntas é como se não tivesse havido "adeus" e regressamos aos "bons velhos tempos".

À minha "mui antiga" colega de casa, já dos tempos de licenciatura na Caparica, Margarida, uma das poucas pessoas que consegue fazer-me chorar de rir, e rir de chorar. Uma grande amiga com uma enorme boa disposição e um coração ainda maior. Que a nossa amizade perdure assim por muitos mais anos.

À minha mais "recente" colega de casa (como um ano passa depressa), Raquel, amiga e quase como uma mãe na casa de Lisboa. Sempre atenciosa e com uma palavra amiga que ameniza e eleva o espírito e moral, temos o repertório já cheio de gargalhadas e histórias partilhadas.

Às minhas colegas e amigas de mestrado, Catarina, Mafalda, Pipa e Susana, obrigada pelos bons momentos dentro e fora de obrigações escolares, pelos passeios inesperados, pelas gargalhadas e histórias partilhadas. Olhando para trás, é muito interessante o que um trabalho de grupo pode fazer, as amizades que daí podem resultar.

Quanto à minha família, é difícil colocar em palavras todo o meu apreço e agradecimento por todo o apoio que me têm dado, não só agora, mas durante toda a minha vida. Não sou uma pessoa que demonstre muito os seus sentimentos, mas quero (e espero) que saibam que os amo muito.

Ao meu avô, uma estrelinha no céu que me guia nas noites mais sombrias, que sempre recordei com tanto carinho e ternura para comigo. Não te tive muito tempo comigo, mas sei que, onde quer que estejas, estás a olhar por mim, e sinto a tua presença ao meu lado. Gosto muito de ti, e nunca te esquecerei.

À minha avó Diamantina, que desde sempre me lembro a chamar-me "magana", com um sorriso maroto como quem quer brincadeira com a neta. Sei que com o passar do tempo, vem o que chamas de "velhice" (como eu digo, "velhos são os trapos"). Mas cada vez que conversamos ao telefone, parece que tudo isso se dissipa, de tanto jovial que se parece a tua voz. Apesar de não te dar todos os beijinhos que queres, o meu amor por ti não muda, eu é

que sou mesmo assim. Ainda assim, quero que saibas que gosto muito de ti, e que vou tentar exprimir mais os meus afectos.

Ao meu avô Carreira, que tem sempre um sorriso no rosto e é um brincalhão por natureza. Sei que os últimos tempos não têm sido fáceis para ti, e consequentemente para mim, mas tens e terás sempre o meu apoio em tudo o que eu conseguir ajudar. Ainda iremos voltar a tocar acordeão (ou eu o pouco que consegui aprender contigo até agora) na casa do campo e passar belos dias tranquilos junto da natureza. Gosto muito de ti, e quero um dia ser tão brincalhona e bem-disposta como tu.

À minha avó Cila, que sorri com os olhos "à chinês" de tão pequeninos que ficam, e me dá vontade de rir. Também para ti não têm sido fáceis os últimos tempos, e (volto a dizer) consequentemente para mim. Contudo, sabes que podes contar comigo para tudo o que eu conseguir fazer para ajudar, e continuaremos a ter muitos "momentos de confessionário" para trocar histórias e sorrisos. Gosto muito de ti, e quero ter um dia a tua paciência e tranquilidade, pois assim a vida é mais agradável.

Aos meus pais, que tanto me têm aturado em todos os sentidos, na teimosia, na ambição, na felicidade e na concretização de sonhos. Tenho tanto para vos dizer, e ainda assim, faltam-me as palavras. Quero ter a tua boa disposição e paciência mãe, e a tua força de vontade e persistência pai. É-me impossível descrever tudo o que sinto por vocês, que sempre foram o meu apoio para tudo, os meus pilares, nos bons e maus momentos. Gosto tanto de vocês que nem imaginam, porque (mais uma vez) não sou muito boa a transmitir os meus sentimentos. Mas sendo assim, ainda bem que tenho esta oportunidade de vos dizer tudo isto (haverá sempre aquele bocadinho que fica por dizer), que gosto, que vos amo muito, e que faria tudo por vocês.

Foi um ano difícil, pelo que queria (e espero) ter(-me) superado para ser uma melhor pessoa e profissional, não melhor do que uma outra pessoa qualquer, mas melhor do que o meu próprio "eu" do passado, do ontem, ambicionando sempre por um melhor futuro do que o esperado.